

SEARCH REQUEST FORM

Requestor's Name: _____ Serial Number: _____
Date: _____ Phone: _____ Art Unit: _____

Search Topic:

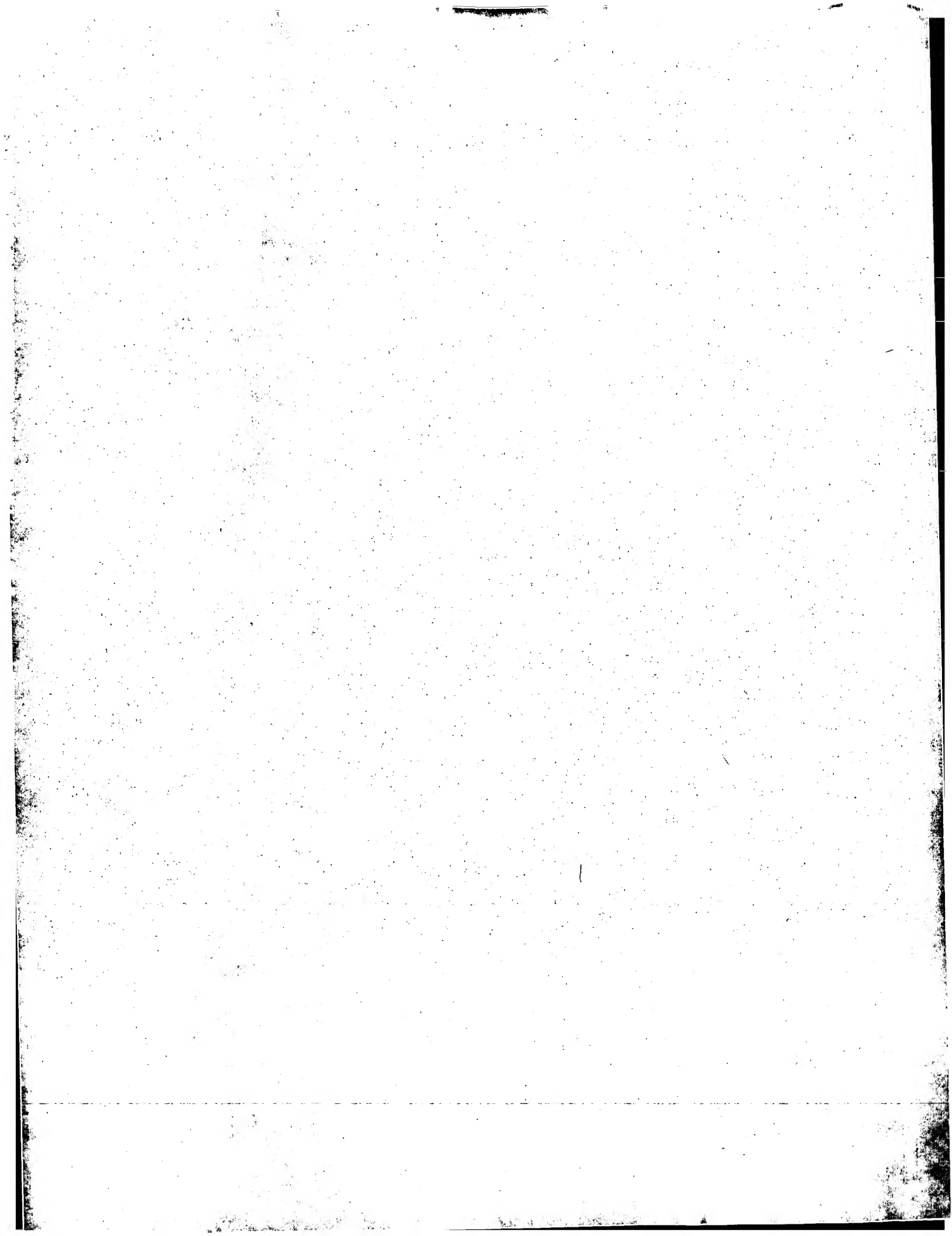
Please write a detailed statement of search topic. Describe specifically as possible the subject matter to be searched. Define any terms that may have a special meaning. Give examples or relevant citations, authors keywords, etc., if known. For sequences, please attach a copy of the sequence. You may include a copy of the broadest and/or most relevant claim(s).

STAFF USE ONLY

Date completed: 01-17-03
Searcher: Beverly e4994
Terminal time: 26
Elapsed time: _____
CPU time: _____
Total time: 28
Number of Searches: 1
Number of Databases: _____

Search Site
____ STIC
____ CM-1
____ Pre-S
Type of Search
____ N.A. Sequence
____ A.A. Sequence
____ Structure
____ Bibliographic

Vendors
____ IG Suite
____ STN
____ Dialog
____ APS
____ Geninfo
____ SDC
____ DARC/Questel
____ Other



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 6615.05 seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-1
Perfect score: 3984
Sequence: .1,accttggtgctggagc.....agagttttgaaatttttt 3984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3984	100.0	3984	6	AR201455	AR201455 Sequence
2	3980.4	99.9	4307	9	AF104413	AF104413 Homo sapi
3	3775	94.8	4241	9	AF164041	AF164041 Homo sapi
4	2483	62.3	3213	6	AR201456	AR201456 Sequence
5	2483	62.3	3213	10	AF104414	AF104414 Mus muscu
6	2422	60.8	2442	6	E24613	E24613 warts Prote
7	1518.2	38.1	80146	9	AL583963	E24613 warts Prote
8	1360.4	34.1	1374	6	E24614	E24614 warts Prote
9	930.8	23.4	1807	17	AF119846	AF119846 Homo sapi
10	848.4	21.3	16490	2	AC102712	AC102712 Mus muscu
11	788.6	19.8	795	6	E24616	E24616 warts Prote
12	760	19.1	3155	6	AR201457	AR201457 Sequence
13	760	19.1	3460	10	AB023958	AB023958 Mus muscu
14	744.4	18.7	5486	6	E38226	E38226 Human tumor
15	744.4	18.7	5486	6	E38227	E38227 Human tumor
16	742.8	18.6	4098	9	AF207547	AF207547 Homo sapi
17	741.8	18.6	3412	9	AB028019	AB028019 Homo sapi
18	680.4	17.1	164490	2	AC102712	AC102712 Mus muscu
19	604.8	15.2	191604	2	AC126128	AC126128 Rattus no
20	580.4	14.6	582	6	E24615	E24615 warts Prote
21	526.6	13.2	181750	2	AC098623	AC098623 Rattus no
22	503.2	12.6	181750	2	AC098623	AC098623 Rattus no
23	470.6	11.8	5360	3	DROWARTS	L39837 Drosophila
24	469	11.8	5720	3	DMU29608	U29608 Drosophila
25	469	11.8	5720	6	AR201454	AR201454 Sequence
26	340	8.5	183597	9	AL356285	AL356285 Human DNA
27	281.2	7.1	1322	3	EOC249683	EOC249683 Euplotes
28	281.2	7.1	1578	3	EOC249680	EOC249680 Euplotes
29	275	6.9	1804	3	EOC249682	EOC249682 Euplotes
30	269.4	6.8	1935	6	AR139101	AR139101 Sequence
31	269.4	6.8	4725	9	BC028603	BC028603 Homo sapi
32	269.4	6.8	4983	6	AX056364	AX056364 Sequence
33	269.4	6.8	5181	9	AB023182	AB023182 Homo sapi
34	269	6.8	2202	8	MCPRTKINA	Z30329 M.crySTALLI
35	255.8	6.4	2422	8	AY128279	AY128279 Arabidops
36	255.6	6.4	1606	3	EOC249681	EOC249681 Euplotes
37	255.4	6.4	1314	3	EOC249684	EOC249684 Euplotes
38	255.4	6.4	1745	3	AF275634	AF275634 Caenorhab
39	254.6	6.4	3018	6	A52140	A52140 Sequence 6
40	254.6	6.4	3018	6	AR084691	AR084691 Sequence
41	254.6	6.4	3018	9	HSPROKINX	Z35102 H.sapiens m
42	254.6	6.4	3593	9	BC012085	BC012085 Homo sapi
43	250.6	6.3	1738	8	AY096444	AY096444 Arabidops
44	250.6	6.3	2365	8	AY063969	AY063969 Arabidops
45	248.6	6.2	1807	3	ECU47679	ECU47679 Euplotes cr

ALIGNMENTS

RESULT 1	AR201455	3984 bp	DNA	linear	PAT 20-APR-2002
LOCUS	AR201455	Sequence 3 from patent US 6359193.			
DEFINITION	AR201455	Sequence 3 from patent US 6359193.			
ACCESSION	AR201455	GI:20252343			
VERSION	AR201455.1				
KEYWORDS	Unknown.				
SOURCE	Unknown.				
ORGANISM	Unknown.				
REFERENCE	1 (bases 1 to 3984)				
AUTHORS	Xu,T., Tao,W., Wang,W., Zhang,S. and Yu,W.				
TITLE	Nucleotide sequences of lats genes				
JOURNAL	Patent: US 6359193-A 3 19-MAR-2002;				
FEATURES	Location/Qualifiers				

source	1..3984																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																																									
--------	---------	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--	--

Db 2041 ATAAAGAAAGAAACAGATTAACAACTTCACCTATTACTGTTTAGGAAAAACAAGAAATG 2100
QY |||||||
Db 2101 AAGACGGAAGGAATCTCGTATTCAAGATTATCTCCCAAGCAATTTAAATTTCTTTATGG 2160
QY |||||||
Db 2101 AAGACGGAAGGAATCTCGTATTCAAGATTATCTCCCAAGCAATTTAAATTTCTTTATGG 2160
QY |||||||
Db 2161 AGCAACATGTAGAAAATGACTCAATCTCATCAGCAGCGTCTACATCGTAAAGAAACAAAT 2220
QY |||||||
Db 2221 TAGAGAATGAATGATCGGGTGTGATTATCTCAAGATGCCAGGATCAATGAGAAAGA 2280
QY |||||||
Db 2221 TAGAGAATGAATGATCGGGTGTGATTATCTCAAGATGCCAGGATCAATGAGAAAGA 2280
QY |||||||
Db 2281 TGCTTTGGCAAAAGAACTTAATTAACATCCGCTTTAAAGGGCTAAATGGACAAGTCTA 2340
QY |||||||
Db 2341 TGCTTTGGCAAAAGAACTTAATTAACATCCGCTTTAAAGGGCTAAATGGACAAGTCTA 2340
QY |||||||
Db 2341 TGCTTTGGCAAAAGAACTTAATTAACATCCGCTTTAAAGGGCTAAATGGACAAGTCTA 2400
QY |||||||
Db 2341 TGCTTTGGCAAAAGAACTTAATTAACATCCGCTTTAAAGGGCTAAATGGACAAGTCTA 2400
QY |||||||
Db 2401 AAGTAGACTAAGGCTTTGTATGCAACAAAACCTCTTCGAAAGAAAGATGTTCTTCCTC 2460
QY |||||||
Db 2401 AAGTAGACTAAGGCTTTGTATGCAACAAAACCTCTTCGAAAGAAAGATGTTCTTCCTC 2460
QY |||||||
Db 2461 GAAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCCTGGCTGAAGCTGACAATGAAT 2520
QY |||||||
Db 2461 GAAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCCTGGCTGAAGCTGACAATGAAT 2520
QY |||||||
Db 2521 GGGTAGTTCGCTATATTATTCATTCACCAAGATAAGGCAATTTATACCTTTGTAATGGACT 2580
QY |||||||
Db 2521 GGGTAGTTCGCTATATTATTCATTCACCAAGATAAGGCAATTTATACCTTTGTAATGGACT 2580
QY |||||||
Db 2581 ACATTCCTGGGGTGATATGATGAGCCCTATTAAATAGAAATGGGCAATCTTCCAGAAAGTC 2640
QY |||||||
Db 2581 ACATTCCTGGGGTGATATGATGAGCCCTATTAAATAGAAATGGGCAATCTTCCAGAAAGTC 2640
QY |||||||
Db 2641 TGGCAGCATCTACATAGCAGAACTTACCTGTCAGTTCGAAAGTCTTCAATAAATGGGTT 2700
QY |||||||
Db 2641 TGGCAGCATCTACATAGCAGAACTTACCTGTCAGTTCGAAAGTCTTCAATAAATGGGTT 2700
QY |||||||
Db 2701 TTATTCATAGAGATATTAAACCTGATAATATTTTTCATGATCGTGTGATGATGATGATGAT 2760
QY |||||||
Db 2701 TTATTCATAGAGATATTAAACCTGATAATATTTTTCATGATCGTGTGATGATGATGATGAT 2760
QY |||||||
Db 2761 TGACTGACTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
QY |||||||
Db 2761 TGACTGACTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2820
QY |||||||
Db 2821 GTGGTACCATCCACGCAAGATAGCATGATTTTCAGTAATGAATGGGGGATCCCTCAA 2880
QY |||||||
Db 2821 GTGGTACCATCCACGCAAGATAGCATGATTTTCAGTAATGAATGGGGGATCCCTCAA 2880
QY |||||||
Db 2881 GCTGTCGATGTGGACAGACTGAAGCCATTAGAGCGGAGAGCTGACGCCAGCAGCAGCAGC 2940
QY |||||||
Db 2881 GCTGTCGATGTGGACAGACTGAAGCCATTAGAGCGGAGAGCTGACGCCAGCAGCAGCAGC 2940
QY |||||||
Db 2941 GATGCTAGACATCTTTGGTGGGACTCCCAATTAATTTGACCTGAAAGTGTGCTAC 3000
QY |||||||
Db 2941 GATGCTAGACATCTTTGGTGGGACTCCCAATTAATTTGACCTGAAAGTGTGCTAC 3000
QY |||||||
Db 3001 GAACAGGATACACAGTTGTTGATGTTGGGAGTGGTGTGTTTCTTTTGAAGTGT 3060
QY |||||||
Db 3001 GAACAGGATACACAGTTGTTGATGTTGGGAGTGGTGTGTTTCTTTTGAAGTGT 3060
QY |||||||
Db 3061 TGGTGGGCAACCTCTTTCTTGGCACAACACACCAATTAAGAAACAAATGAAGTTATCA 3120
QY |||||||
Db 3061 TGGTGGGCAACCTCTTTCTTGGCACAACACCAATTAAGAAACAAATGAAGTTATCA 3120
QY |||||||
Db 3121 ACTGCAACATCTCTTCCATTTCCACCAAGCTTAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3180
QY |||||||

Db 3121 ACTGCAACATCTCTTCCATTTCCACATTTCCACCACAAGCTAACTCAGTCTGAGCTTCTGATC 3180
QY |||||||
Db 3181 TTATTATTAAACTTTTGGCGAGGACCGGAAGATCGCTTAGGCAAGATGCTGCTGATGAA 3240
QY |||||||
Db 3181 TTATTATTAAACTTTTGGCGAGGACCGGAAGATCGCTTAGGCAAGATGCTGCTGATGAA 3240
QY |||||||
Db 3241 TAAAGCTCATCCATTTTAAACAAATTTGACTTCTCCAGTACCTTGACAGCAGCAGCTG 3300
QY |||||||
Db 3241 TAAAGCTCATCCATTTTAAACAAATTTGACTTCTCCAGTACCTTGACAGCAGCAGCTG 3300
QY |||||||
Db 3301 CTTTCATACATTTCTTAAATCACACACCCCAACAGATACATCAATTTTGTGCTGTTGATC 3360
QY |||||||
Db 3301 CTTTCATACATTTCTTAAATCACACACCCCAACAGATACATCAATTTTGTGCTGTTGATC 3360
QY |||||||
Db 3361 CTGATAAATATTAGGAGTGATGATAAGGAGAGAAATGTAATGACACTCTCAATGGAT 3420
QY |||||||
Db 3361 CTGATAAATATTAGGAGTGATGATAAGGAGAGAAATGTAATGACACTCTCAATGGAT 3420
QY |||||||
Db 3421 GGTATAAATGGAAGCATCTCTGAACATGCTATCTTATGAATTTACCTTTCCGAAAGTCTT 3480
QY |||||||
Db 3421 GGTATAAATGGAAGCATCTCTGAACATGCTATCTTATGAATTTACCTTTCCGAAAGTCTT 3480
QY |||||||
Db 3481 TTGATGACAAATGGCTACCCATATAATATCCGAGCCTATTGAATGATGATGATGATGAT 3540
QY |||||||
Db 3481 TTGATGACAAATGGCTACCCATATAATATCCGAGCCTATTGAATGATGATGATGATGAT 3540
QY |||||||
Db 3541 CACAAGGCTCAGCAGCAGCTCGGATGAGATGATCAAAACACAGCTCAGAGATTAATA 3600
QY |||||||
Db 3541 CACAAGGCTCAGCAGCAGCTCGGATGAGATGATCAAAACACAGCTCAGAGATTAATA 3600
QY |||||||
Db 3601 ATCGCGATCTAGTATATGTTTAAACACACTAGTAAATGTAATGAGGATTTGTAAGAG 3660
QY |||||||
Db 3601 ATCGCGATCTAGTATATGTTTAAACACACTAGTAAATGTAATGAGGATTTGTAAGAG 3660
QY |||||||
Db 3661 GGCCTGAAATGGAGGCTGTTTTCAGGCTCAGAGTAAATGTAATGAGGATTTGTAAGAG 3720
QY |||||||
Db 3661 GGCCTGAAATGGAGGCTGTTTTCAGGCTCAGAGTAAATGTAATGAGGATTTGTAAGAG 3720
QY |||||||
Db 3721 TATATGTCCTCTGTCGATCAATATTTTATTTTCTTAAATGTAATGAGGATTTGTAAGAG 3780
QY |||||||
Db 3721 TATATGTCCTCTGTCGATCAATATTTTATTTTCTTAAATGTAATGAGGATTTGTAAGAG 3780
QY |||||||
Db 3781 AATGTTAATTTTCCAGCGCTTAAATCAGTATTTAGAAAAAATTTGTTAAGGAAAG 3840
QY |||||||
Db 3781 AATGTTAATTTTCCAGCGCTTAAATCAGTATTTAGAAAAAATTTGTTAAGGAAAG 3840
QY |||||||
Db 3841 TAAATTTGAACCTGAATATTATAGTCACTTCTGTTTAAAGTAAAGTAAAGTAAAGTAAAG 3900
QY |||||||
Db 3841 TAAATTTGAACCTGAATATTATAGTCACTTCTGTTTAAAGTAAAGTAAAGTAAAGTAAAG 3900
QY |||||||
Db 3901 TGCTTTGTTTAAAGGAGAAACCTGCTATCTTATTTGTTATATGCTAAATAATTTTAA 3960
QY |||||||
Db 3901 TGCTTTGTTTAAAGGAGAAACCTGCTATCTTATTTGTTATATGCTAAATAATTTTAA 3960
QY |||||||
Db 3961 TACAAGAGTTTGTGAAATTTTTT 3984
QY |||||||
Db 3961 TACAAGAGTTTGTGAAATTTTTT 3984
QY |||||||

RESULT 2

AF104413

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

AF104413 Homo sapiens large tumor suppressor 1 (LATS1) mRNA, complete cds.
AF104413 Homo sapiens large tumor suppressor 1 (LATS1) mRNA, complete cds.
AF104413.1 GI:4324433

Homo sapiens

Homo sapiens

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

1 (bases 1 to 4307)

Xu, T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.

Identifying tumor suppressors in genetic mosaics: the Drosophila

Qy	1203	GT	TGGCAGACAACC	AATCATCATG	CAGAGTTC	TACCAAA	TTTA	CTTCAT	CAGGAGA	1262		
Db	1521	GT	TGGCAGACAACC	AATCATCATG	CAGAGTTC	TACCAAA	TTTA	CTTCAT	CAGGAGA	1580		
Qy	1263	CT	TGGAATGCAGA	ATGGTACT	GCAGCAAA	CTGATT	TTCTAT	GATAC	CAACCAAA	TGTTGCCCT	1322	
Db	1581	CT	TGGAATGCAGA	ATGGTACT	GCAGCAAA	CTGATT	TTCTAT	GATAC	CAACCAAA	TGTTGCCCT	1640	
Qy	1323	GC	TGGCACTGTGA	ATCGGACG	CAACCA	CACTCT	CTGAC	AGAGTAA	TGSGACAA	1382		
Db	1641	GCT	GGCACTGTGA	ATCGGACG	CAACCA	CACTCT	CTGAC	AGAGTAA	TGSGACAA	1700		
Qy	1383	AG	CCCTCTCTG	TTTACAA	ACAGGGG	ATCTG	CTGCT	CTTCTCAT	TATACA	ATGGAAGT	1442	
Db	1701	AG	CCCTCTCTG	TTTACAA	ACAGGGG	ATCTG	CTGCT	CTTCTCAT	TATACA	ATGGAAGT	1760	
Qy	1443	ATT	CCTCAGTCTAT	GATGGT	GCCAA	CAAGAA	ATAGT	CATA	ACATGGA	CTATATAACAT	1502	
Db	1761	ATT	CCTCAGTCTAT	GATGGT	GCCAA	CAAGAA	ATAGT	CATA	ACATGGA	CTATATAACAT	1820	
Qy	1503	AG	TGTACTTGG	ACTGCA	AAACAA	ATTTGG	CGCTCAGT	CACTTCTG	CTCCAG	CGCCAGTCA	1562	
Db	1821	AG	TGTACTTGG	ACTGCA	AAACAA	ATTTGG	CGCTCAGT	CACTTCTG	CTCCAG	CGCCAGTCA	1880	
Qy	1563	CC	GAGAGTGGG	CAATAA	TCCTCAT	NGGCA	ACCTAC	ATACAGT	GAGTCA	AAATTC	1622	
Db	1881	CC	GAGAGTGGG	CAATAA	TCCTCAT	NGGCA	ACCTAC	ATACAGT	GAGTCA	AAATTC	1940	
Qy	1623	TT	TATAACCC	ATTAGG	AAATAG	ACAAGT	CACTCTG	CTAA	TTCTCAG	CCCTTCTGCTACA	1682	
Db	1941	TT	TATAACCC	ATTAGG	AAATAG	ACAAGT	CACTCTG	CTAA	TTCTCAG	CCCTTCTGCTACA	2000	
Qy	1683	AC	AGTCACTG	CAATAC	ACAG	CTCTTAT	TAACAG	CGCTG	GAAGAT	TGCGGTATTA	1742	
Db	2001	AC	AGTCACTG	CAATAC	ACAG	CTCTTAT	TAACAG	CGCTG	GAAGAT	TGCGGTATTA	2060	
Qy	1743	AA	ACCAGAG	CTACAG	ACTG	CTTTAG	CACTAC	ACAC	ACCCTCT	CTG	ATACACAG	1802
Db	2061	AA	ACCAGAG	CTACAG	ACTG	CTTTAG	CACTAC	ACAC	ACCCTCT	CTG	ATACACAG	2120
Qy	1803	CAA	ACTGTTCA	ACCAG	TCCTTTT	CTGAGG	AAACCG	TCCTCAA	ATGTG	ACTGTGATGCCA	1862	
Db	2121	CAA	ACTGTTCA	ACCAG	TCCTTTT	CTGAGG	AAACCG	TCCTCAA	ATGTG	ACTGTGATGCCA	2180	
Qy	1863	CT	TGTTGCTGA	AGCTCC	AAACTAT	CAAGG	ACAC	CAAC	CCCTAC	CCCAAA	1922	
Db	2181	CT	TGTTGCTGA	AGCTCC	AAACTAT	CAAGG	ACAC	CAAC	CCCTAC	CCCAAA	2240	

Qy	2283	CTTTGCCAAAAGAACTTAATTAACATCCGCTCTTTAAAGGGCTAAAAATGGCAAGACTCTATG	2342
Db	2601	CTTTGCCAAAAGAACTTAATTAACATCCGCTCTTTAAAGGGCTAAAAATGGCAAGACTCTATG	2660
Qy	2343	TTTCTGGAAGATAAAGACACATAGGAATAGAGACATTTGGTGAAGTCTGTCTAGCAAGAAA	2402
Db	2661	TTTCTGGAAGATAAAGACACATAGGAATAGAGACATTTGGTGAAGTCTGTCTAGCAAGAAA	2720
Qy	2403	GTAGATACTAAGGCTTTGTATGCAACAAAACCTCTTGCAAAAGAAAGATGTTCTTCTTCGA	2462
Db	2721	GTAGATACTAAGGCTTTGTATGCAACAAAACCTCTTGCAAAAGAAAGATGTTCTTCTTCGA	2780
Qy	2463	AATCAAGTCGCTCATGTTTAAGCGTAGAGAGATATCCCTGGCTGGAAGCTGCAATGAATGG	2522
Db	2781	AATCAAGTCGCTCATGTTTAAGCGTAGAGAGATATCCCTGGCTGGAAGCTGCAATGAATGG	2840
Qy	2523	GTAGTTCGCTCATATTAATTAATTCATCCCAAGATAAGGACAATTTATACTTTGTAATGGACTAC	2582
Db	2841	GTAGTTCGCTCATATTAATTCATCCCAAGATAAGGACAATTTATACTTTGTAATGGACTAC	2900
Qy	2583	ATTCCCTGGGGTGATATGATGACGCTTAATTAATAGAATGGGCATCTTTCCAGAAAGCTCG	2642
Db	2901	ATTCCCTGGGGTGATATGATGACGCTTAATTAATAGAATGGGCATCTTTCCAGAAAGCTCG	2960
Qy	2643	GCAGATTCTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTCATPAAAATGGGTTTT	2702
Db	2961	GCAGATTCTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTCATPAAAATGGGTTTT	3020
Qy	2703	ATTCTATAGAGATATTAACCTGATATATTTTGAATGTAGCTGATGTCATATTAATATG	2762
Db	3021	ATTCTATAGAGATATTAACCTGATATATTTTGAATGTAGCTGATGTCATATTAATATG	3080
Qy	2763	ACTGACTTTGGCCTCTGCACCTGGCTTCAGATGGACACACGATTTCTTAAGTACTATCAGAGT	2822
Db	3081	ACTGACTTTGGCCTCTGCACCTGGCTTCAGATGGACACACGATTTCTTAAGTACTATCAGAGT	3140
Qy	2823	GGTGACCATCCAGCGCAAGATAGCATGGATTTCACTAATGAATGGGGGATCCCTCAAGC	2882
Db	3141	GGTGACCATCCAGCGCAAGATAGCATGGATTTCACTAATGAATGGGGGATCCCTCAAGC	3200
Qy	2883	TGTCGATGTGGAGACAGACTGAAGCCATTAAGAGCGGAGAGCTGCACGCCACACACGCA	2942
Db	3201	TGTCGATGTGGAGACAGACTGAAGCCATTAAGAGCGGAGAGCTGCACGCCACACACGCA	3260
Qy	2943	TGCTAGCACATCTTTTGGTTGGGACTCCCAATTAATTTGCACCTGAAGTCTTGTCTACGA	3002
Db	3261	TGCTAGCACATCTTTTGGTTGGGACTCCCAATTAATTTGCACCTGAAGTCTTGTCTACGA	3320
Qy	3003	ACAGGATACACAGTTGTGTGATTTGGTGGAGTGTGGTGTATTTCTTTTGAAGTGTG	3062
Db	3321	ACAGGATACACAGTTGTGTGATTTGGTGGAGTGTGGTGTATTTCTTTTGAAGTGTG	3380
Qy	3063	GTGGGCAACCTCTCTTTCTTGGCACAACACACATTTAGAAACACAAATGAAGTTTATCAAC	3122
Db	3381	GTGGGCAACCTCTCTTTCTTGGCACAACACACATTTAGAAACACAAATGAAGTTTATCAAC	3440
Qy	3123	TGGCAAAACATCTCTTCACATTTCCACACAAGCTAAACCTCAGTCCCTGAAGCTTCTGATCCT	3182
Db	3441	TGGCAAAACATCTCTTCACATTTCCACACAAGCTAAACCTCAGTCCCTGAAGCTTCTGATCCT	3500
Qy	3183	ATTATTAACATTTTGGCGAGGACCCGAAGATTCGCTTTAGGCAAGAATGGTGTGATGAATA	3242
Db	3501	ATTATTAACATTTTGGCGAGGACCCGAAGATTCGCTTTAGGCAAGAATGGTGTGATGAATA	3560
Qy	3243	AAAGCTCATCCATTTTTTTAAACAATTGACTTCTCCAGTGACCTGACGACAGCAGCTGTCT	3302
Db	3561	AAAGCTCATCCATTTTTTTAAACAATTGACTTCTCCAGTGACCTGACGACAGCAGCTGTCT	3620
Qy	3303	TCATACATCTCCTAAAATTCACACACCCACAGATACATCAAAATTTTGATCTGTGTTGATCCT	3362
Db	3621	TCATACATCTCCTAAAATTCACACACCCACAGATACATCAAAATTTTGATCTGTGTTGATCCT	3680
Qy	3363	GATAAATTTATGGAGTGATGATAACGAGGAAGAAAATGTAAATGACACTCTCAATGGGATGG	3422

```

|||||
Db 3681 GATAAATATGAGTATGATATAACGAGGAGAAATGTAATGACACTCTCAATGATGG 3740
Qy 3423 TATATAATGGAAGCATCTGACATGATCTATGAAATTTACCTTCGGAAGTTTTTT 3482
Db 3741 TATAAATGGAAGCATCTGACATGATCTATGAAATTTACCTTCGGAAGTTTTTT 3800
Qy 3483 GATGACAAATGCTACCATATAATATCCGAAGCCTATGAAATGATGAAATATCA 3542
Db 3801 GATGACAAATGCTACCATATAATATCCGAAGCCTATGAAATGATGAAATATCA 3860
Qy 3543 CAAGGCTCAGACGAGCAGTGGATGAGATGATCAACACACAGCCTCAGAGATTA 3602
Db 3861 CAAGGCTCAGACGAGCAGTGGATGAGATGATCAACACACAGCCTCAGAGATTA 3920
Qy 3603 CGCATCTAGTATATGTTTAAACACACAGTAAATGTAATGAGGATTTGTAAGGG 3662
Db 3921 CGCATCTAGTATATGTTTAAACACACAGTAAATGTAATGAGGATTTGTAAGGG 3980
Qy 3663 CCTGAAATGCGAGGTGTTTGAAGTCTCGAGAGTAAATATGCAATATGACAGACTA 3722
Db 3981 CCTGAAATGCGAGGTGTTTGAAGTCTCGAGAGTAAATATGCAATATGACAGACTA 4040
Qy 3723 TATATGTCGTCTGTGTACATATTTTCTTCTTAAATATGCGAAATCCTTTTAAA 3782
Db 4041 TATATGTCGTCTGTGTACATATTTTCTTCTTAAATATGCGAAATCCTTTTAAA 4100
Qy 3783 TGTAAATTTATCCAGCGCTTTAAATCAGTATTTAGAAAAAAATTTGTAAGGAA 3842
Db 4101 TGTAAATTTATCCAGCGCTTTAAATCAGTATTTAGAAAAAAATTTGTAAGGAA 4160
Qy 3843 AATTATGACATGAATATATATGATGATCTTCTGTTTAAATGATGATGATGATG 3902
Db 4161 AATTATGACATGAATATATATGATGATCTTCTGTTTAAATGATGATGATGATG 4220
Qy 3903 CTTTCTTAAAGGAGAAACCTGGTATCTATTGTTATATGCTAAATATTTTAAATA 3962
Db 4221 CTTTCTTAAAGGAGAAACCTGGTATCTATTGTTATATGCTAAATATTTTAAATA 4280
Qy 3963 CAAGAGTTTTGAAATTTTTTT 3984
Db 4281 CAAGAGTTTTGAAATTTTTTT 4302

RESULT 3
AF164041 4241 bp mRNA linear PRI 06-DEC-1999
LOCUS Homo sapiens WARTS protein kinase (WARTS) mRNA, complete cds.
DEFINITION AF164041
ACCESSION AF164041
VERSION AF164041.1 GI:5738135
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1. (bases 1 to 4241)
Kishimoto,T., Niva,S.-i., Nagamine,Y., Nishiyama,Y. and Saya,H.
WARTS protein, polynucleotide encoding the same, antisense
polynucleotide thereof, and antibody recognizing the protein
Patent: Japan (PCT/JP98/03739) 24-AUG-1998;
Sumitomo Electric Industries, Ltd.; 1 Taya-cho, Sakae-ku,
Yokohama-shi; Kanagawa;
Japan;
2. (bases 1 to 4241)
Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Yamamoto,H., Hiroka,T., Kitamura,N. and
Saya,H.
A human homolog of Drosophila warts tumor suppressor, h-warts,
localized to mitotic apparatus and specifically phosphorylated
during mitosis
FEBS Lett. 459 (2), 159-165 (1999)
JOURNAL 9947636
MEDLINE 10518011
PUBMED

```

```

3. (bases 1 to 4241)
Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Nakamura,H., Koga,H. and Saya,H.
Direct Submission
Submitted (30-JUN-1999) Tumor Genetics and Biology, Kumamoto
University School of Medicine, 2-2-1 Honjo, Kumamoto 860-0811,
Japan
Location/Qualifiers
1. 4241
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. 4241
/gene="WARTS"
33..3425
/gene="WARTS"
/notes="similar to Drosophila warts tumor suppressor;
localized on mitotic apparatus, specifically
phosphorylated at mitotic checkpoint"
/codon_start=1
/product="WARTS protein kinase"
/protein_id="AAD50272.1"
/db_xref="GI:5738136"
/translation="MKRSEKPEGYPMRPTKTPASNYTVSSROMLOEIRSLRLNSKP
SDAAKAEHMSKMTEDPQVRNPPKFTGTHKALOEIENSLIPANETNSRSSTSEVN
PQMLQDLQAGFEDMVIQALQKTNRRSIEAIEFISKMSQDPRREQMAAARIPIN
ASMKPGNVQOSNRKQSWKSKESLVPQHPGLGESVAYHSESPNSQTDVGRPLSGS
GISAFVQAHPSNGRVNPPPPQVRSVTPPPRGTTPPPRGTTPPPSWEPNSQTKR
YSGNMEYVLSRISPVPGAWOBYGPPPLNTSPMPPNQOGQISSVPVQSPILIMOS
SKSNFPPSGRCMONGTQOTDFMIHONVVPAGTVNRPVPLTAANGQSPALQTG
GSAAPSSYITNGIPQSMVMVNRNSHIMELYNISVPLQTNMPOSSSAPASPSGHE
IPTQNPPIVRSENPNLGRNSANSQPSATTVTATITAPIQPVKSMRLKPEL
QTALAPTHPSWIPIQTVTPSPFPGTASNVTVMPVAENYQPPPPPKHLLHQ
NPSPVPEYISKSPKEDQSLPKEDSEKSEYENDSGDKKKQITTSPTITKNNKKDE
ERRESIQSYSPQAFKFFMEQVENVLKSQORHRKQLENMRVLSQDDQADOMR
KMLCKESNYIRLAKRMDKSMFKVITKTLGAFGEVCLARKVDYKALYATKLRKDK
VILRNOYHVAERDILAEADNEWVRLYYSFQDKNLYFVMDYIPGGDMMSLLRMG
IPESLAREYIAELTCAVESVHKMGFTIHRDKPNLIDROGHIKLTDFGLCTGRVTP
HDSKYVSGDHPQSDMSFNQWDFSSCRCDRLKPLERRAAROHORHLSHVGTP
NTIAPVLELTGTQCLDMSVGVILFEMLVGQPPFLAQTPLETOMKVINWQTSHP
POAKLSPEASDLIIKLCRGPEDRLKNGADEKAHPFFKTIIDFSSDROOSASYIPKI
THPTDTSNFPDVPDPDKLWSDNEENNDLNGWYKNGKHPEHAFYEFTFRFFDDING
YPVNPKEIYEYVINSQSEQSDDDQNTSGEIKNRDLVYV"
misc_feature
2145..3059
/gene="WARTS"
/notes="Region: serine/threonine kinase domain"
BASE COUNT 1384 a 884 c 817 g 1156 t
ORIGIN

```

```

Query Match 94.8%; Score 3775; DB 9; Length 4241;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3786; Conservative 0; Mismatches 0; Indels 1; Gaps 1;
Qy 199 CCTTCGTGTGGCTACATATATAGATGTTTTCATGAGAGAGTGAAAGCCAGAGAT 258
Db 1 CCTTCGTGTGGCTACATATATAGATGTTTTCATGAGAGAGTGAAAGCCAGAGAT 60
Qy 259 ATAGACAAATGAGCGCTAAGACCTTTCTGCCAGTACTACTGTCAGTAGCGGCAAA 318
Db 61 ATAGACAAATGAGCGCTAAGACCTTTCTGCCAGTACTACTGTCAGTAGCGGCAAA 120
Qy 319 TGTTCACAGAAATTCGGGAATCCCTTATAGAAATTTATCTAAACCATCTGCTCTAAG 378
Db 121 TGTTCACAGAAATTCGGGAATCCCTTATAGAAATTTATCTAAACCATCTGCTCTAAG 180
Qy 379 CTGACATAACATGATGATAAATGTCAACCGAAGATCTCGACAGTCAAGAATCCACCA 438
Db 181 CTGACATAACATGATGATAAATGTCAACCGAAGATCTCGACAGTCAAGAATCCACCA 240
Qy 439 AATTGGGACCATCATAAAGCCTTTCAGGAAATTCGAAACTCTCTGCTCCATTTGCAA 498
Db 241 AATTGGGACCATCATAAAGCCTTTCAGGAAATTCGAAACTCTCTGCTCCATTTGCAA 300
Qy 499 ATGAACAAATTTCTTCTCGGAGTACTTTCAGAGTTAATCCACAAATGCTTCAAGACTTGC 558

```


Db 301 ATGAACAAATCTTCTCGAGTACTTCAGAGTTAATCCACAAATGCTTCAAGACTTGC 360
 QY 559 AAGCTGCTGGATTTCATGAGGATGGTTATACAGCTCTTCAGAAAGTCTTCAACACAGCA 618
 Db 361 AAGCTGCTGGATTTCATGAGGATGGTTATACAGCTCTTCAGAAAGTCTTCAACACAGCA 420
 QY 619 GTATAGAGCAGCAATGAATTCATTAGTAAATGAGTTACCAAGATCCTCGACGAGC 678
 Db 421 GTATAGAGCAGCAATGAATTCATTAGTAAATGAGTTACCAAGATCCTCGACGAGC 480
 QY 679 AGATGGCTGCAGCAGCTCCAGACCTATTAATGCGAGCATGAAACAGGAGTGTGCAGC 738
 Db 481 AGATGGCTGCAGCAGCTCCAGACCTATTAATGCGAGCATGAAACAGGAGTGTGCAGC 540
 QY 739 AATCAGTTTAACCGCAACAGAGCTGGAAGGTTCTTAAAGATCCTTACTTCTCAGAGC 798
 Db 541 AATCAGTTTAACCGCAACAGAGCTGGAAGGTTCTTAAAGATCCTTACTTCTCAGAGC 600
 QY 799 ATGCCGCCCTACTAGGAGAAAGTGTGGCTATCATTTCTGAGAGTCCCAACTCACAGACAG 858
 Db 601 ATGCCGCCCTACTAGGAGAAAGTGTGGCTATCATTTCTGAGAGTCCCAACTCACAGACAG 660
 QY 859 ATGAGGAGACCTTTGTCTGGATCTGTATATCAGCATTTTGTTCAGCTCACCCTTAGCA 918
 Db 661 ATGAGGAGACCTTTGTCTGGATCTGTATATCAGCATTTTGTTCAGCTCACCCTTAGCA 720
 QY 919 ACGACACAGAGTAGTGAACCCCTCCAGAGGTACAACTCCACCTCCCTTCTCAGGAGC 978
 Db 721 ACGACACAGAGTAGTGAACCCCTCCAGAGGTACAACTCCACCTCCCTTCTCAGGAGC 780
 QY 979 CTCGAAGAGGCGAGCTCCCTCCAGAGGTACAACTCCACCTCCCTTCTCAGGAGC 1038
 Db 781 CTCGAAGAGGCGAGCTCCCTCCAGAGGTACAACTCCACCTCCCTTCTCAGGAGC 840
 QY 1039 CAACCTCTCAACAAAGCGCTATTTCTGGAACATGGAATAGTATCTCCGGAATCTCTC 1098
 Db 841 CAACCTCTCAACAAAGCGCTATTTCTGGAACATGGAATAGTATCTCCGGAATCTCTC 900
 QY 1099 CTGTCCCACTGGGCGTGGCAAGAGGCTATCTCCACCACTCTCAACACTTCCCTCCCA 1158
 Db 901 CTGTCCCACTGGGCGTGGCAAGAGGCTATCTCCACCACTCTCAACACTTCCCTCCCA 960
 QY 1159 TGAATCTCTTAATCAAGGACAGAGGCGATTAGTCTCTGTTCTGTTGGCAGACAACAA 1218
 Db 961 TGAATCTCTTAATCAAGGACAGAGGCGATTAGTCTCTGTTCTGTTGGCAGACAACAA 1020
 QY 1219 TCATCATGACAGTTCTAGCAAAATTAATCTTCCATCAGGAGACCTGGAATGCAAGATG 1278
 Db 1021 TCATCATGACAGTTCTAGCAAAATTAATCTTCCATCAGGAGACCTGGAATGCAAGATG 1080
 QY 1279 GTACTGGCAAACTGATTTATGATACACCAAAATGTTGCTCCCTGCTGGCACTGTGAATC 1338
 Db 1081 GTACTGGCAAACTGATTTATGATACACCAAAATGTTGCTCCCTGCTGGCACTGTGAATC 1140
 QY 1339 GCGAGCCACCACTCCATATCTCTGACAGCAGCTTAATGGACAAGCCCTTCTGCTTTAC 1398
 Db 1141 GCGAGCCACCACTCCATATCTCTGACAGCAGCTTAATGGACAAGCCCTTCTGCTTTAC 1200
 QY 1399 AAACAGGGGATCTGCTGCTCTGCTCATATACAAATGGAAGTATTCCTCAGCTTAGA 1458
 Db 1201 AAACAGGGGATCTGCTGCTCTGCTCATATACAAATGGAAGTATTCCTCAGCTTAGA 1260
 QY 1459 TGTTGCCAAACAGAAATAGTACAACTGGAACATATATACATAGTCTACCTGACCTGC 1518
 Db 1261 TGTTGCCAAACAGAAATAGTACAACTGGAACATATATACATAGTCTACCTGACCTGC 1320
 QY 1519 AAACAAATTTGCCCTCAGCTATCTTCTGCTCCAGCCAGCTATCCCGAGCAGTGGGATG 1578
 Db 1321 AAACAAATTTGCCCTCAGCTATCTTCTGCTCCAGCCAGCTATCCCGAGCAGTGGGATG 1380
 QY 1579 AAATCCCTTACATGGCAACCTTAACATACAGTGAAGTCAAAATCTCTTTTAATACCACTTAG 1638

Db 1381 AAATCCCTTACATGGCAACCTTAACATACCACTGAGCTCAAAATCTTTTAATACCACTTAG 1440
 QY 1639 GAAATAGAGCAGTCACTCTCTTAATCTCAGCTTCTGCTACACAGCTCACTCAATTA 1698
 Db 1441 GAAATAGAGCAGTCACTCTCTTAATCTCAGCTTCTGCTACACAGCTCACTCAATTA 1500
 QY 1699 CACCAGCTCTCTTATCAACAGAGCTGTGAAAGTATGCTGTATTAACACAGAGCTTACAGA 1758
 Db 1501 CACCAGCTCTTATCAACAGAGCTGTGAAAGTATGCTGTATTAACACAGAGCTTACAGA 1560
 QY 1759 CTGCTTTAGCAGCTTACACACCTTCTTGATACACAGCTTCAAACTGTTCACCA 1818
 Db 1561 CTGCTTTAGCAGCTTACACACCTTCTTGATACACAGCTTCAAACTGTTCACCA 1620
 QY 1819 GTCCCTTTCTCAGGAGACCTTCAAACTGTACCTGTGATGCCACCTGTTCGAGCTC 1878
 Db 1621 GTCCCTTTCTCAGGAGACCTTCAAACTGTGATGCCACCTGTTCGAGCTC 1680
 QY 1879 CAACTATCAAGGACACCACTTACCCAAACATCTGTGTCACCAAAACCTCATCTG 1938
 Db 1681 CAACTATCAAGGACACCACTTACCCAAACATCTGTGTCACCAAAACCTCATCTG 1740
 QY 1939 TTCCTCCATAGAGTCAATCAGTAAAGCTAGCAAGAGGATCAGCAAGCTTGCCTCAGG 1998
 Db 1741 TTCCTCCATAGAGTCAATCAGTAAAGCTAGCAAGAGGATCAGCAAGCTTGCCTCAGG 1800
 QY 1999 AAGATGAGAGTGAAGAGTATGAAATGTGTAGTGGGATGAAGAAAGAAACAGA 2058
 Db 1801 AAGATGAGAGTGAAGAGTATGAAATGTGTAGTGGGATGAAGAAAGAAACAGA 1860
 QY 2059 TTACAACCTTCACTTACTTGTAGGAAACCAAGAAAGATCAAGAGCGAAGGGAATCTC 2118
 Db 1861 TTACAACCTTCACTTACTTGTAGGAAACCAAGAAAGATCAAGAGCGAAGGGAATCTC 1920
 QY 2119 GTATTCAGAGTTATCTCCTCAAGCATTTAAATCTTTATGGAGCAACATGTAGAAAATG 2178
 Db 1921 GTATTCAGAGTTATCTCCTCAAGCATTTAAATCTTTATGGAGCAACATGTAGAAAATG 1980
 QY 2179 TACTCAAACTCTCATCAGCAGCTCTACATCGTAAAAACAATAGAGAAATCAATGATGC 2238
 Db 1981 TACTCAAACTCTCATCAGCAGCTCTACATCGTAAAAACAATAGAGAAATCAATGATGC 2040
 QY 2239 GGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAAGAAAT 2298
 Db 2041 GGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAAGAAAT 2100
 QY 2299 CTAATTCATCCGCTTTAAAAAGGCTTAAATGGCAAGCTATGTTTGTGAGATFAAAGA 2358
 Db 2101 CTAATTCATCCGCTTTAAAAAGGCTTAAATGGCAAGCTATGTTTGTGAGATFAAAGA 2160
 QY 2359 CACTAGGAATAGGAGCATTTGCTGAAGTCTCTAGCAAGAAAGTACTAGTACTAGGCTT 2418
 Db 2161 CACTAGGAATAGGAGCATTTGCTGAAGTCTCTAGCAAGAAAGTACTAGTACTAGGCTT 2220
 QY 2419 TGTATGCAACAAACCTCTTCGAAGAAAGATGTTTCTTCGAAATCAAGTCTGCTCATG 2478
 Db 2221 TGTATGCAACAAACCTCTTCGAAGAAAGATGTTTCTTCGAAATCAAGTCTGCTCATG 2280
 QY 2479 TTAAGGCTGAGAGATATCTGCTGAAGCTGCAATGAATGGTAGTCTGCTCATATTT 2538
 Db 2281 TTAAGGCTGAGAGATATCTGCTGAAGCTGCAATGAATGGTAGTCTGCTCATATTT 2340
 QY 2539 ATTCAATCCAAAGTAAAGGACAAATTTATCTTGTAAATGGACTACATCTCTGGGCTGATA 2598
 Db 2341 ATTCAATCCAAAGTAAAGGACAAATTTATCTTGTAAATGGACTACATCTCTGGGCTGATA 2400
 QY 2599 TGATGAGCTTATTAATTAAGATGGCACTTTCGAGAAAGTCTGCGACGATTTCTACATAG 2658
 Db 2401 TGATGAGCTTATTAATTAAGATGGCACTTTCGAGAAAGTCTGCGACGATTTCTACATAG 2460
 QY 2659 CAGAACTTACCTGTCAGTGTGAAAGTGTTCATAAAATGGGTTTTTATTCATAGAGATATTA 2718
 Db 2461 CAGAACTTACCTGTCAGTGTGAAAGTGTTCATAAAATGGGTTTTTATTCATAGAGATATTA 2520

QY 2719 AACCTGATAATATTTGATGATCGTGATGGTCAATATAAATGACTGACTTTGGCCTCT 2778
 Db 2521 AACCTGATAATATTTGATGATCGTGATGGTCAATATAAATGACTGACTTTGGCCTCT 2580
 QY 2779 GCACTGGCTTCAGATGGACACAGATCTAAGTACTATCAGAGTGGTACCATCCAGGC 2838
 Db 2581 GCACTGGCTTCAGATGGACACAGATCTAAGTACTATCAGAGTGGTACCATCCAGGC 2640
 QY 2839 AAGATGACATGGATTTTCACTAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGACA 2898
 Db 2641 AAGATGACATGGATTTTCACTAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGACA 2700
 QY 2899 GACTGAAGCCATTAGACGGGAGCTGCACGGCAGACAGAGATGTCTAGACATCTTT 2958
 Db 2701 GACTGAAGCCATTAGACGGGAGCTGCACGGCAGACAGAGATGTCTAGACATCTTT 2760
 QY 2959 TGGTTGGGACTCCCAATATATTCACCTGAAGTGTGTGCTACGAACAGGATACACAGT 3018
 Db 2761 TGGTTGGGACTCCCAATATATTCACCTGAAGTGTGTGCTACGAACAGGATACACAGT 2820
 QY 3019 TGTCTGATTTGGTGGAGTGTGGTGTATTTTGTGAATGTTGGTGGGACAACTCCCTT 3078
 Db 2821 TGTCTGATTTGGTGGAGTGTGGTGTATTTTGTGAATGTTGGTGGGACAACTCCCTT 2880
 QY 3079 TCTTGGCACAACACCAATTAGAACAACAATGAAGTTATCACTGCGCAACATCTCTTC 3138
 Db 2881 TCTTGGCACAACACCAATTAGAACAACAATGAAGTTATCACTGCGCAACATCTCTTC 2940
 QY 3139 ACATTCCACCACAAGCTTAAGCTCAGTCTGCTGAGCTTCTGATCTTATTTAACTTTGCC 3198
 Db 2941 ACATTCCACCACAAGCTTAAGCTCAGTCTGCTGAGCTTCTGATCTTATTTAACTTTGCC 3000
 QY 3199 GAGACCCGAAGATCGGTTAGGCAAGAATGGTGTGATGAATAAAGCTCATCCATTTT 3258
 Db 3001 GAGACCCGAAGATCGGTTAGGCAAGAATGGTGTGATGAATAAAGCTCATCCATTTT 3060
 QY 3259 TTAACAACATTTGCTTCCAGTCACTGACAGACAGCTGCTGCTTATACATTTCTTAAAA 3318
 Db 3061 TTAACAACATTTGCTTCCAGTCACTGACAGACAGCTGCTGCTTATACATTTCTTAAAA 3120
 QY 3319 TCACACACCCACAGATACATCAATTTTGTATCTGTGATCTCTGATGAATAATATGAGAGT 3378
 Db 3121 TCACACACCCACAGATACATCAATTTTGTATCTGTGATCTCTGATGAATAATATGAGAGT 3180
 QY 3379 ATGATACGGAGAGAAATGTAATGACACTCTCAATGGATGCTATAAATAAGTGAAGC 3438
 Db 3181 ATGATACGGAGAGAAATGTAATGACACTCTCAATGGATGCTATAAATAAGTGAAGC 3240
 QY 3439 ATCTTGAACATGCAATCTATGAATTTACCTTCCGAAGGTTTTTTTGTGACAAATGGCTACC 3498
 Db 3241 ATCTTGAACATGCAATCTATGAATTTACCTTCCGAAGGTTTTTTTGTGACAAATGGCTACC 3300
 QY 3499 CATATAATATCCGAAGCCTATTGAATATGAATATCAATTAATTCAGAGGCTCAGAGAGC 3558
 Db 3301 CATATAATATCCGAAGCCTATTGAATATGAATATCAATTAATTCAGAGGCTCAGAGAGC 3360
 QY 3559 AGTCGGATGAAGATGATCAAAACACAGCTCAGAGATTAAAAATCGGATCTAGTATATG 3618
 Db 3361 AGTCGGATGAAGATGATCAAAACACAGCTCAGAGATTAAAAATCGGATCTAGTATATG 3420
 QY 3619 TTTAACACACTAGTAAATTAATGATGAGGATTTGTAAGGGCCTGAAATGGAGGTG 3678
 Db 3421 TTTAACACACTAGTAAATTAATGATGAGGATTTGTAAGGGCCTGAAATGGAGGTG 3480
 QY 3679 TTTTGAAGTCTGAGAGTAAATTAATGCAATATGACAGAGCTATATATGTTGCTCTGT 3738
 Db 3481 TTTTGAAGTCTGAGAGTAAATTAATGCAATATGACAGAGCTATATATGTTGCTCTGT 3540
 QY 3739 GTACAATATTTTATTTCTTAAATTTATGGGAAATCCTTTTAAATGTTTAAATTTT -CCA 3797
 Db 3541 GTACAATATTTTATTTCTTAAATTTATGGGAAATCCTTTTAAATGTTTAAATTTTATTCOCA 3600

QY 3798 GCGTTTAAATCAGTATTTAGAAAAAATTTGTATAAGGAAAGTAAATATGAACTGAAT 3857
 Db 3601 GCGTTTAAATCAGTATTTAGAAAAAATTTGTATAAGGAAAGTAAATATGAACTGAAT 3660
 QY 3858 ATTATAGTCAGTCTTGTGGTACTTAAAGTACTTAAATAAGTACTGTTTGTGTTAAAGGA 3917
 Db 3661 ATTATAGTCAGTCTTGTGGTACTTAAAGTACTTAAATAAGTACTGTTTGTGTTAAAGGA 3720
 QY 3918 GAACTGTGATCTATTTGTATATATGCTAAATAATTTTAAATAACAGAGTTTTTGA 3977
 Db 3721 GAACTGTGATCTATTTGTATATATGCTAAATAATTTTAAATAACAGAGTTTTTGA 3780
 QY 3978 TTTTGT 3984
 Db 3781 TTTTGT 3787
 RESULT 4
 AR201456 LOCUS AR201456 3213 bp DNA linear PAT 20-APR-2002
 DEFINITION Sequence 5 from patent US 6359193.
 ACCESSION AR201456
 VERSION AR201456.1 GI:20252344
 KEYWORDS
 SOURCE Unknown.
 ORGANISM Unknown.
 REFERENCE 1 (bases 1 to 3213)
 AUTHORS Xu, T., Rao, W., Wang, W., Zhang, S. and Yu, W.
 TITLE Nucleotide sequences of late genes
 JOURNAL Patent: US 6359193-A 5 19-MAR-2002;
 FEATURES
 Location/Qualifiers
 1..3213
 source /organism="unknown"
 BASE COUNT 946 a 764 c 696 g 807 t
 ORIGIN
 Query Match 62.3%; Score 2483; DB 6; Length 3213;
 Best Local Similarity 87.4%; Pred. No. 0;
 Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;
 QY 732 GTGCAGCAATCAGTTTACCCGCAACAGAGCTGGAAGGTTCTAAAGATCCTTAGTTCT 791
 Db 1 GTGCAGCAATCAGTTTACCCGCAACAGAGCTGGAAGGTTCTAAAGATCCTTAGTTCT 60
 QY 792 CAGAGGATGCGCCGCGCAGCTAGGAGAAAGTGTGCGCTATCATCTCTGAGAGTCCCACTCA 851
 Db 61 CAGAGACACGCGCCATCTCTAGGAGAAATGTGTTTATCGTTCTGAAAGCCCACTCA 120
 QY 852 CAGACAGATGAGAGACCTTTGTCTGGATCTGGTATATCAGCATTTGTTCAAGCTCAC 911
 Db 121 CAGGCGATGAGAGACCTCTGTCTGGATCGGCAATTCAGCATTTGTTCAAGCTCAC 180
 QY 912 CCTAGCAACGACAGAGAGTGAACCCGCCACCACTCAAGTAAAGAGGTGTTACTCTCT 971
 Db 181 CCAAGCAATGACAGAGAGTGAACCCGCCACCACTCAAGTAAAGAGGTGTTACTCTCT 240
 QY 972 CCACCACTCCAGAGAGCCAGACTCCCGCTCCAGAGGTACACTCCACCTCCCGCTCA 1031
 Db 241 CCACCACTCCAGAGAGCCAGACTCCCGCTCCAGAGGTACACTCCCGCTCCCGCTCA 300
 QY 1032 TGGCAACCAACTCTCAACAAAGCGCTATTCTGGAACATGGAATACGTAATCTCCGA 1091
 Db 301 TGGCAACCAACTCTCAACAAAGCGCTATTCTGGAACATGGAATACGTAATCTCCGA 360
 QY 1092 ATCTCTCTTCCACCTCGGGGATGGAAGAGGGCTATCTCCACACCTCTCAACACT 1151
 Db 361 ATCTCTCTTCCACCTCGGGGATGGAAGAGGGTACCTCCACACCTCTTACCCT 420
 QY 1152 TCCCGCATGATCTCTCTATCAAGAGGCAATAGTCTCTCTCTCTCTCTCTCTCTCTCTCT 1211
 Db 421 TCTCCCATGATCT 480

Qy	1212	CAACCAATCATCATGACAGATTCTTAGCAAAATTTAACTTTCCATCAGGAGACACTGGGAATG	1217
Db	481	CAACCCATCATCATGACAGAGTACTAGCAAAATTTAACTTTACACAGGCGCACCTGGAGTT	540
Qy	1272	CAGAATGGTATGCGACAAACTCATTTCAATGATACACCAAAATGTTGTCCCTGCTGGCACT	1331
Db	541	CAGAAATGGTGGTGCAGTCTCATTTATATCGTGGCACCAAAA--TGTCGCCCATCTGGTCTCT	597
Qy	1332	GTGAATCGGCAGCCACCACTCCATATCTCTGCAGCAGCACTAATGACCAAAAGCCCTTCT	1391
Db	598	GTCACTCGGCAGCCACCACTCCATATCTCTGCAGCCAGCTAATGCAAAAGCCCTCT	657
Qy	1392	GCCTTTAAACAAGGGGATCTGCTGCTCTGCTCATATACAAATGGAAGTATFTCTCTCAG	1451
Db	658	GCCTTTAAACAAGGGGCTTCTGCTGCTCCACCATATCGGCCAATGGAACCTGCTCTCAG	717
Qy	1452	TCATATGATGTCGCAACAAGAAATAGTCATACATACATGGAACATATATACATTAAGTACCT	1511
Db	718	TCGATGATGGTCCCAACAGGAACAGTCATACATATGGAGCTTTTATAATATTAATGTCCCT	777
Qy	1512	GGACTGCAACAAATTTGGCCTCAGTCATCTTCTGCTCAGGCCCACTCATCCCCGAGCAGT	1571
Db	778	GGACTGCAAAAGCCTGGCCCCAGTCGCTTCTGCTCTCGCAGTCATCCCCAAGCGGT	837
Qy	1572	GGGCATGAATCCCTACATGCGCAACTAACATACACAGTGAGGTCAAAATCTTTTAATAAC	1631
Db	838	GGGCATGAATTCCTACATGGCAAGCTAACATACACAGTGAGGTCAAAATCTTTTAATAAC	897
Qy	1632	CCATTAGGAATAGACGAAGTCACCTGCTTAATCTCAGCCTTCTGCTACACAGTCACCT	1691
Db	898	CCATTAGGAAGTAGACAAGTCACCTGCTTAATCTCAGCCTTCTGCTACACAGTCACCT	957
Qy	1692	GCATTTACACCAGCTCCTATTCCACAGCCTGTGAAAAGTAGTCGTGTATTAAAACACAGAG	1751
Db	958	GCATCATCACCCGCTCTATTCAACAGCCGCTGAAAAGCATGCGCTCTCTGAACACAGAG	1017
Qy	1752	CTACAGACTGCTTTTAGCACTACACACCTCTTTGGATACACAGCAATTCACAACTGTT	1811
Db	1018	CTGCAGACTGCTTTAGCCCCAACCCATCTCTTTGGATGCCACAGCCAGTTCAGACTGTT	1077
Qy	1812	CAACCCAGTCCCTTTTCCCTGAGGGAACCGCTTCAAAATGTGACTGTGATGCCACCTGTTGCT	1871
Db	1078	CAGCCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCTCCACCCTGTTGCT	1137
Qy	1872	GAAGCTCCAAACTATCAAGGACCCACCACTACCCAAACAAATCTGCTGCACCAAAAC	1931
Db	1138	GAAGCTCCAAAGTATCAGGTCACCAACCGCTTATCCAAACAAATCTGCTACACCAAAAC	1197
Qy	1932	CCATCTGTTCTCCATACAGTCAATCAGTAAGCCTAGCAAGAGGATCAGCCAAAGCTTG	1991
Db	1198	CCATCTGTTCTCCATATCAGTCAGTAAAGTAAAGCCTGCAAAAGTAAACAGCTTAGCTTA	1257
Qy	1992	CCCAAGGAGATGAGAGTCAAAAGAGTTATGAAAATGTTGATAGTGGGGATAAAGAAAAG	2051
Db	1258	CCCAAGGAGATGATAGTGAAGAAGTCCGACAGTGGTGACTCTGGGGATAAAGAAAAG	1317
Qy	2052	AAACAGATTACAACCTTCACTATTACTGTTAGGAAAAACAAGAAAGATGAAGAGCGAAGG	2111
Db	1318	AAACAGATTACAACCTTCACTATTACTGTTAGGAAAAACAAGAAAGATGAAGAACGAGA	1377
Qy	2112	GAATCTGTTTCAAAGTTATTCTCTCAAGCATTTAAATCTTTATGGAGCAACATGTA	2171
Db	1378	GAGTCTCGGATTCAGATTACTTCCCAACAGGCTTTTAAGTCTTTCATGGAGCAGCAGTA	1437
Qy	2172	GAAAATGTACTCAATCTCATCAGCAGGCTTACATCGTAAACAAATTAGAGAAATGAA	2231
Db	1438	GAGAAGCTCTTGAAGTCTCATCAGCAGGCTTGTGATCGGAAGAAAGCAGCTAGAAAAATGAA	1497
Qy	2232	ATGATCGGGTGGATTATCTCAAGATGCCAGCATCAAAATGAGAAAGATGCTTTGCCAA	2291
Db	1498	ATGATCGGGTGGATTATCTCAAGATGCCAGCATCAAAATGAGAAAGATGCTTTGCCAG	1557
Qy	2292	AAAGAACTTAATTACATCCGCTTTAAAAGGGCTAAAATGGCAAGCTATGTTGTTGGAAG	2351

[illegible]

Congress Ave BCM 254D, New Haven, CT 06536, USA

Location/Qualifiers

1. .3213

/organism="Mus musculus"

/db_xref="taxon:10090"

/tissue_type="brain"

/dev_stage="newborn"

<1. .3213

/gene="Lats1"

<1. .2889

/gene="Lats1"

/note="putative serine threonine kinase; (MUM) Lats1"

/codon_start=1

/product="large tumor suppressor 1"

/protein_id="AADI6883.1"

/db_xref="GI:4324436"

/translation="VQHSINRKQSWKSGKESLPQRHGPSLGRGVNVVRSFSPNSQADV
GRSLSSGIAFAQAPNSGCRVNPVPPVRSVTPPPRGOTTTPPPPPSME
PSSQTKRYSNMETVTSRISFVPVPAQGEYGPPLTSPNPPSOAQRALSSVPQGSF
QPTIMQSTKFNTPPGVGGQGSDFIVHONVPGSVTPROPPPYTPVANGQSF
SALQTGASAAPPFANGVPSQSMVNNRNSHNMELYINVPGLQTANPQSSSAPQAQSS
PSCGHELTQPNITPVRSNFNNPLGSRASHANSQPSATVTAITPAIPQQPVMMSS
VLKPELTALATPHSNMPPQVTPPTFSEGTASSPVVPIVAAEAPVQPPPPYTV
KHLLHONPSVRYESVSKPCKDEQSLPKDEKDSKSHOORLHRKKOLENEMRVGLSQQ
KNKDERRERISYQSPQAFKFMEOCHVNLVSKHOOHLRKGTCVCLARKVDTKALYATV
AQDQMRMLCKGKSNYTLRAKMDKSMVKIKTLGTCFCEVCLARKVDTKALYATV
TLRKKQVLLRNQVAHVAEREDILAEADNENVYRLYYSPQDKNLYFVMDYIPGGDMMS
LLIRMGFVLPENLAFYIAETCAVSEVHKMGFTHRDIKPNILLDRGHKLTDPGLG
TGFTWTHDSYVQSGDHPQSDMSFNSWEGDPSPNCRGDRLLKPLERRAAQHOHQRLA
SLVGTPNVIAPEVLLRTGYQLCDWNPQSVILCEMLYQGPFFLAQTPLETKQKVLV
TSLHVPQAKLSPEASDILIKLCRGPEDRLGKNGADEIKAHPPFKTIDFSSDLRQOQS
SYIPKTHPTDTSNEDVPDPLKWSGDGSEENISDTLSGYNKNGKHEPAHYEFTFR
FFDNGQPVNPKPIEYVTHSGQSGDEDDQHTSDGNNRLVYV"

764. 596. 807

0.6

FEATURES

source

gene

CDS

Query Match	62.3%;	Score	2483;	DB	10;	Length	3213;
Best Local Similarity	87.4%;	Pred.	No. 0;				
Mismatches	0;	Mismatches	385;	Indels	15;	Gaps	
Matches	2769;	Conservative					

QY	732	GTGCAGCAATCAGTTAAACCGCAACACAGACTGGAAGGTTCTAAAGAATCCTTAGTTCCT	791
DB	.1	GTGCACAATTCATTAAACCGAAAACAAGCTGGAAAGTTCTTAAGAGTCTCTAGTTCCT	60
QY	792	CAGAGGCATGGCCGCCCATCTAGSAGAAAGTGCGCTATCATCTTGAGAGTCCCACACTCA	851
DB	61	CAGACACACGGCCCCATCTCTAGSAGAAATGTGGTTTTATCTGTTCTGAAAGCCCAACTCA	120
QY	852	CAGACAGATGTAGGAAGACCTTTGTCTGCATCTGGTATATCAGCATTTCTCAAGCTCAC	911
DB	121	CAGGCGGATGTAGGAAGACCTCTGTCTGATCGGCAITTCGACATTTGCTCAAGCTCAC	180
QY	912	CCTAGCAACGGACAGAGCTGAACCCCCACCACACCTCAAGTAAGGAGTGTTACTCCT	971
DB	181	CCAAGCAANTGGACAGAGGTGAACCCCCACCACACCTCAAGTTAGGAGTGTTACTCCT	240
QY	972	CCACCACCTCCAAGAGGCCAGACTCCCCCTCCCAAGAGGTACAACCTCCACTCCCCCTTCA	1031
DB	241	CCACCACCTCCGAGAGGCCAGACCCCACTCCCGGAGGCCACCACTCCCTCCGCCCTCA	300
QY	1032	TGGGAACCAAACCTCTCAACAAAGCCGTATTCTGGAACATGGAATAGTAAATCTCCCGA	1091
DB	301	TGGGAACCAAGCTCTCAGAACAAACGCGTACTCTGGGAACATGGAGTAGTACGTACCTCCGA	360
QY	1092	ATCTCTCTGTGCCACTGGGGCATGCAAGAGGGGTATCCTCCACACCTCTCAACACT	1151
DB	361	ATCTCCCTGTTCACCTGGGGCTGGCAGAGGGGTACCTCCACCACTCTTACCACCT	420
QY	1152	TCCCCCATGAATCCTCCTTAATCAAGACAGAGGCCATTAGTTCGTCTGTGGGCGAGA	1211
DB	421	TCCTCCATGAATCCCCCTAGCCAGGCTCAGAGGGCCATTAGTTCGTTCAGATTGGTAGA	480
QY	1212	CAACCAATCATGACAGAGTTCTAGCAANTTTAACTTTCCATCAGGAGACCTTGGGAATG	1271

Db	1558	AAAGAGTCTAACTATATTCGTCTTTAAAGGGCTAAATATGACAAAGTCTATGTTTGTAAAG	1611
QY	2352	ATAAGACACTAGGAATAGCAGCATTTGGTGAAAGTCTGTCTAGCAAGAAAGTAGATACT	2411
Db	1618	ATAAGACATTAGGAATAGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTCGATACT	1677
QY	2412	AAGCTTTTGATGCAACAAAAACTCTTTCGAAAGAAAGATGTTCTTCTTCGAAATCAAGTC	2471
Db	1678	AAAGCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGGTG	1737
QY	2472	GCTCATGTTAAGCTGAGAGAGATATCCTTGGCTGAAGCTGACAATGAATGGGTAGTTCGT	2531
Db	1738	GCTCATGTGAAGCGGAGAGGGATATCCTTASGAGAAGCCACAATGAGTGGTGGTCCCG	1797
QY	2532	CTATATTATTCAITTCCAAGATAAGGACAAATTTATACTTTCTAATGGACTACATTCCTGGG	2591
Db	1798	CTGTACTACTCTTTCAGGACAGGACAACCTTGACTTTGTGATGGACTACATTCCTGGG	1857
QY	2592	GGTGATATGATGAGCCTATTAAATAGAAATGGGCATCTTCCAGAAAAGTCTGGCACGATTC	2651
Db	1858	GGGSAATATGATGAGCCTATTAAATAGAAATGGGCATCTTCTCGAAAATCTGGCACGATTC	1917
QY	2652	TACATAGCAACAACCTTACCTGTGCAGTTGAAGTGTTCATATAAATGGSTTTTATTCATAGA	2711
Db	1918	TACATAGCAACAACCTTACCTGTGCAGTTGAAGTGTTCATATAAATGGSTTTTATTCATAGA	1977
QY	2712	GATATTAAACCTTGATATAATTTTTGATTGATCGTGATGGTGCATATTAAATGACGTGACTTT	2771
Db	1978	GATATTAAACCTTGATATAATTTTTGATTGATCGTGATGGTGCATATTAAATGACGTGACTTT	2037
QY	2772	GGCTCTGCACTGGCTTCAGATGGACACACGATTTAAAGTACTATCAGAGTGTGACCAT	2831
Db	2038	GGCTTGTGCACCTGGCTTCAGATGGACACATGACTCCAAGTACTACCAGAGTGGGATCAC	2097
QY	2832	CCAGGCAAGATAGCATGGATTTTCAGTAATGATGGGGGATCCCTCAAGCTGTGCATGT	2891
Db	2098	CCAGGCAAGATAGCATGGATTTTCAGTAATGATGGGGGATCCCTCCAATTTCTCGGTGT	2157
QY	2892	GGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCAGCACACGAGTCTCTAGCA	2951
Db	2158	GGGACAGACTGAAGCCACTGGAGCGGAGAGCTGCTGCCAGCACACGAGTCTCTAGCC	2217
QY	2952	CATTCTTTGGTGGGACTCCCAATTATTTGCACCTGAAAGTGTGCTACGAAACAGGATAC	3011
Db	2218	CATTCTCTGGTGGGACTCCCAATTATTTGCACCTGAAAGTGTGCTACGAAACAGGATAT	2277
QY	3012	ACACAGTTGTGATTTGGTGGAGTGTGGTGTATTTCCTTTTGAATGTTGGTGGACAA	3071
Db	2278	ACACAGTGTGTGACTTGGTGGAGTGTGGTGTATTTCCTTTTGAATGTTGGTGGACAA	2337
QY	3072	CCTCTCTTCTTGGCACAACACCATTAGAACAACAATGAAGTGTATCAACTGGCAACA	3131
Db	2338	CCTCTCTTCTTGGCACAACACCCATTAGAAAACAACAATGAAGTGTATCTCTGGCAACT	2397
QY	3132	TCCTTTCACATTCACCCACAAGCTAAACTCAGTCCCTGAAGCTTCTGATCTTATTATAA	3191
Db	2398	TCCTTACACATCCCTCCTCAAGCTAAGCTGAGTCTTGAAGCCCTCTGACCTCAATTACAA	2457
QY	3192	CTTTGGCGAGGCCCGAAGATCCCTTAGGCAAGAATGGTGCTGATGAATAATAAAGCTCAT	3251
Db	2458	CTGTGCGAGGCCGAAGACCCGCCCTGGCAGAAACGGTCTGATGAGATAAAGGCTCAT	2517
QY	3252	CCATTTTTTAAAAAATTTGACTTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATACATT	3311
Db	2518	CCATTTTTTAAAGACCATCGAATTTCTCTAGTGAATCTGAGACAGCAGTCTGCTTCATACATC	2577
QY	3312	CCTTAAATCACACACCCACAGATACATCAAAATTTTGTGCTCTGATCTGATGATAAATTA	3371
Db	2578	CCTTAAATCAGCATCCACAGATACATCCAATTTCCAGCCCTGTTGATCCTGATAAATTC	2637
QY	3372	TGGAGTGATGATAACGAGGAGAAAATGTAAATGACACTCTCAATGGATGGTATATAAAAT	3431
Db	2638	TGAGCGGATGGCAGCGAGGAGAAAATATCAGTGACACTCTGAGCGGATGGTATAAAAAT	2697

QY 3432 GGAAGCATCTGACATGCTATGATTTACCTTCGGAAGGTTTGTGATGACAAT 3491
 Db 2698 GGAAGCACCAGCAGCGCTTTCTATGATTTACCTTCGGAAGGTTTGTGATGACAAT 2757
 QY 3492 GGTACCCATATATATCCGAAGCCTATTGAATATGAATACATTAATTCACAGGCTCA 3551
 Db 2758 GGTACCCATATATATCCGAAGCCTATTGAATATGAATACATTAATTCACAGGCTCA 2817
 QY 3552 GAGCAGCGTGGATGAAGATGATCAAAACACAGGCTCAGAGATTAATAATCCGATCTA 3611
 Db 2818 GAACAACAGTCTGATGAAGATGATCAAAACACAGGCTCAGAGATTAATAATCCGATCTA 2877
 QY 3612 GTATATGTTTAACACACTAGTAATAATGTAATGAGGATTTGTAAGAAGGCTGAATG 3671
 Db 2878 GTGATGTTTAATAAATAGAGATCAATTTGTAAGGCTGAATG 2933
 QY 3672 CGAGGTGTTTGAAGTCTGAGAGTAAATATGCAATATGACAGAGCTATATATGTT 3731
 Db 2934 CAGGGGTTTGAAGTCTGAGAGTAAATATGCAATATGACAGAGCTATATATGTT 2985
 QY 3732 GCTCTGTGATCAATATTTTCCATAATATGGAATATGGAATATGGAATATGGAATATGGAAT 3791
 Db 2986 GCTCTGTGATCAATATTTTCCATAATATGGAATATGGAATATGGAATATGGAATATGGAAT 3045
 QY 3792 ATTCAGCGCTTTAAATCAGTATTTAGAAAAAATTTGTAAGAAGGCTGAATATGGAAT 3851
 Db 3046 ATTCAGCGCTTTAAATCAGTATTTAGAAAAAATTTGTAAGAAGGCTGAATATGGAAT 3105
 QY 3852 CTGAATATATGATGAGTCTTTGATGATTAAGTACTTAAATGAATGAATG 3900
 Db 3106 CTGATATATGATGAGTCTTTGATGATTAAGTACTTAAATGAATGAATG 3154

RESULT 6
 E24613
 Locus
 DEFINITION 2442 bp DNA linear PAT 18-JUN-2001
 warts protein, polynucleotide encoding the protein, antisense
 polynucleotide thereof and antibody recognizing the protein.
 E24613
 ACCESSION E24613.1 GI:13024646
 VERSION JP 1999089580-A/1.
 KEYWORDS
 SOURCE unidentified.
 ORGANISM unclassified.
 REFERENCE 1: (bases 1 to 2442)
 AUTHORS Toshiko, K., Shinichi, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
 TITLE polynucleotide encoding the protein, antisense
 polynucleotide thereof and antibody recognizing the protein
 JOURNAL Patent: JP 1999089580-A 1 06-APR-1999;
 SUMITOMO ELECTRIC IND LTD
 COMMENT OS Unidentified
 PN JP 1999089580-A/1
 PD 06-APR-1999
 PE 24-SEP-1997 JP 1997258689
 PR TOSHIOKI KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
 YASUYUKI NISHIYAMA,
 PI HIDEYUKI SAVA
 PC C12N15/09,C07K14/435,C07K16/18,C12Q1/68,G01N33/53, PC
 G01N33/532/C12P21/02,
 PC (C12N15/09,C12R1:91),(C12P21/02,C12R1:19),C12N15/00,
 (C12N15/00, PC C12R1:91)
 CC Strandedness: Double;
 CC Topology: Linear;
 FH Key Location/Qualifiers
 FT Source 1..2442 /organism="Unidentified"
 FT Location/Qualifiers
 1..2442 /organism="unidentified"
 /db_xref="taxon:32644"
 BASE COUNT 777 a 528 c 491 g 646 t

ORIGIN

Query Match 60.8%; Score 2422; DB 6; Length 2442;
 Best Local Similarity 99.8%; Pred No. 0;
 Matches 2436; Conservative 0; Mismatches 5; Indels 1; Gaps 1;
 QY 1062 TCTGGAACATGGAATACGTAATCTCCCGAATCTCTCTGTCCTCCACCTGGGCGATGGCAA 1121
 Db 1 TCTGGAACATGGAATACGTAATCTCCCGAATCTCTCTGTCCTCCACCTGGGCGATGGCAA 60
 QY 1122 GAGGGCTATCTCCACACCTCTCAACACTTCCCGCATGAATCTCCCTCAATCAAGCACAG 1181
 Db 61 GAGGGCTATCTCCACACCTTTTCAACACTTCCCGCATGAATCTCCCTCAATCAAGCACAG 120
 QY 1182 AGAGGCATTAGTCTGCTGTTGGCAGACCAACCAATCATCATGCGAGAGTCTTACGAAA 1241
 Db 121 AGAGGCATTAGTCTGCTGTTGGCAGACCAACCAATCATCATGCGAGAGTCTTACGAAA 180
 QY 1242 TTTAACTTTCCATCAGGGAGACCTGGAATGCGAATGGTACTGGACAAACTGATTTCTATG 1301
 Db 181 TTTAACTTTCCATCAGGGAGACCTGGAATGCGAATGGTACTGGACAAACTGATTTCTATG 240
 QY 1302 ATACACCAAAATCTTCCCTGCTGGCAGCTGTGAATGGCAGCCACCACTCCATATCTCT 1361
 Db 241 ATACACCAAAATCTTCCCTGCTGGCAGCTGTGAATGGCAGCCACCACTCCATATCTCT 300
 QY 1362 CTGACAGCAGCTAATGGACAAAGCCCTTCTGTTTACAAACAGGGGATCTGCTGCTCT 1421
 Db 301 CTGACAGCAGCTAATGGACAAAGCCCTTCTGTTTACAAACAGGGGATCTGCTGCTCTCT 360
 QY 1422 TCGTCATATACAAATGGAAGTATTTCTCAGTCTATGATGTCGCAACAGAAATAGTTCAT 1481
 Db 361 TCGTCATATACAAATGGAAGTATTTCTCAGTCTATGATGTCGCAACAGAAATAGTTCAT 420
 QY 1482 AACATGGAACATATATAACATTAGTACTGCTGACCTGCAACAAATTTGGCTCAGTCTCT 1541
 Db 421 AACATGGAACATATATAACATTAGTACTGCTGACCTGCAACAAATTTGGCTCAGTCTCT 480
 QY 1542 TCTGCTCCAGCCAGTCTATCCCGAGCAGTGGGATGAATCCCTACATGCGCAACCTAAC 1601
 Db 481 TCTGCTCCAGCCAGTCTATCCCGAGCAGTGGGATGAATCCCTACATGCGCAACCTAAC 540
 QY 1602 ATACCACTGAGGTCAAAATCTTTTAAACCCATTAGAAATAGAGCAAGTCACTCTGCT 1661
 Db 541 ATACCACTGAGGTCAAAATCTTTTAAACCCATTAGAAATAGAGCAAGTCACTCTGCT 600
 QY 1662 AATTCTCAGCCTTCTGCTACAAACAGTCACTGCAATTCACACAGCTCTCTATTCACAGCCT 1721
 Db 601 AATTCTCAGCCTTCTGCTACAAACAGTCACTGCAATTCACACAGCTCTCTATTCACAGCCT 660
 QY 1722 GTGAAAAGTATGCTGATGCCACCTGTTGCTGAAGCTCCAACTATCAAGGACCAACCA 1781
 Db 661 GTGAAAAGTATGCTGATGCCACCTGTTGCTGAAGCTCCAACTATCAAGGACCAACCA 720
 QY 1782 TCTTGGATACCAAGCAATTTCAAACTGTTCAACCCAGTCTCTTCTGAGGGAACCGCT 1841
 Db 721 TCTTGGATACCAAGCAATTTCAAACTGTTCAACCCAGTCTCTTCTGAGGGAACCGCT 780
 QY 1842 TCAATGTGATGATGCCACCTGTTGCTGAAGCTCCAACTATCAAGGACCAACCA 1901
 Db 781 TCAATGTGATGATGCCACCTGTTGCTGAAGCTCCAACTATCAAGGACCAACCA 840
 QY 1902 CCCTACCAAAACATCTGCTGACCAAAACCCATCTGCTCTCCATAGGAGTCAATCAGT 1961
 Db 841 CCCTACCAAAACATCTGCTGACCAAAACCCATCTGCTCTCCATAGGAGTCAATCAGT 900
 QY 1962 AAGCCTAGCAAAAGGATCAGCCAGCTTGGCCAAAGGAAGATGAGTGAAGAGGTTAT 2021
 Db 901 AAGCCTAGCAAAAGGATCAGCCAGCTTGGCCAAAGGAAGATGAGTGAAGAGGTTAT 960
 QY 2022 GAAATGTTGATGTTGGGATTAAGAAAGACAGATTACAACTTACCTTACTTACTGTT 2081
 Db 961 GAAATGTTGATGTTGGGATTAAGAAAGACAGATTACAACTTACCTTACTTACTGTT 1020

QY 2082 AGGAAAAACAAGAGATGAAGAGCGAAGGGAATCTCGTATTCAAAGTTATTCTCTCAA 2141
DB 1021 AGGAAAAACAAGAGATGAAGAGCGAAGGGAATCTCGTATTCAAAGTTATTCTCTCAA 1080
QY 2142 GCATTAAATCTTTATGGAGCAACATGTAGAAAATGACTCAAAATCTCATCAGCAGGT 2201
DB 1081 GCATTAAATCTTTATGGAGCAACATGTAGAAAATGACTCAAAATCTCATCAGCAGGT 1140
QY 2202 CTACATCTFAAAAAACAATAGAGAAATGAATGATGCGGGTGGGATTTCTCAAGATGCC 2261
DB 1141 CTACATCTFAAAAAACAATAGAGAAATGAATGATGCGGGTGGGATTTCTCAAGATGCC 1200
QY 2262 CAGGATCAATAGAGAAAGATGCTTTGCCAAAAAGAAATCTAATACATCCGCTCTTAAAGG 2321
DB 1201 CAGGATCAATAGAGAAAGATGCTTTGCCAAAAAGAAATCTAATACATCCGCTCTTAAAGG 1260
QY 2322 GCTAAATGGCAAGTCTATGTTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGT 2381
DB 1261 GCTAAATGGCAAGTCTATGTTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGT 1320
QY 2382 GAAGTCTCTAGCAAGAAAGTACTACTAGGCTTTGTATGCAACAAAACACTCTTGA 2441
DB 1321 GAAGTCTCTAGCAAGAAAGTACTACTAGGCTTTGTATGCAACAAAACACTCTTGA 1380
QY 2442 AAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTG 2501
DB 1381 AAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTG 1440
QY 2502 GCTGAAGTGAACAATGAATGGGTAGTTCGTCTATATTTATTCATTCCAAGATAAGGACAA 2561
DB 1441 GCTGAAGTGAACAATGAATGGGTAGTTCGTCTATATTTATTCATTCCAAGATAAGGACAA 1500
QY 2562 TTATACCTTTGAATGGAGTACATTCCTGGGGTGATATGATGAGCCTATTATTAAGATG 2621
DB 1501 TTATACCTTTGAATGGAGTACATTCCTGGGGTGATATGATGAGCCTATTATTAAGATG 1560
QY 2622 GGCATCTTTCCAGAAAGTCTGCGACGATTTACATAGCAGAACTTACCTGTGCAGTTGAA 2681
DB 1561 GGCATCTTTCCAGAAAGTCTGCGACGATTTACATAGCAGAACTTACCTGTGCAGTTGAA 1620
QY 2682 AGTGTTCATAAATGGGTTTTATTTCATAGATATTAACCTGTGATATTTGATGAT 2741
DB 1621 AGTGTTCATAAATGGGTTTTATTTCATAGATATTAACCTGTGATATTTGATGAT 1680
QY 2742 CGTATGTGATATTAATTAATGACTGACCTTGCCTCTGCCTGCTGATGAGACAC 2801
DB 1681 CGTATGTGATATTAATTAATGACTGACCTTGCCTCTGCCTGCTGATGAGACAC 1740
QY 2802 GATTCTAAGTACTATCAGAGTGGTGACCATCCAGCGCAAGATAGCATGATTTCAAGTAAT 2861
DB 1741 GATTCTAAGTACTATCAGAGTGGTGACCATCCAGCGCAAGATAGCATGATTTCAAGTAAT 1800
QY 2862 GAATGGGGGATCCCTCAAGCTGTGATGAGACAGACTGAAGCCATTAAGAGCGGAGA 2921
DB 1801 GAATGGGGGATCCCTCAAGCTGTGATGAGACAGACTGAAGCCATTAAGAGCGGAGA 1860
QY 2922 GCTGACCCGACGACGAGTGTACACATTTCTTGGTGGGACTCCCAATTTATTT 2981
DB 1861 GCTGACCCGACGACGAGTGTACACATTTCTTGGTGGGACTCCCAATTTATTT 1920
QY 2982 GCACCTGAAGTGTCTACGAAACAGGATACACAGTTGTGTGATTTGGTGGAGTCTTGGT 3041
DB 1921 GCACCTGAAGTGTCTACGAAACAGGATACACAGTTGTGTGATTTGGTGGAGTCTTGGT 1980
QY 3042 GTTATCTTTTGAATGTTGGTGGGCAACCTCTCTTTGGGCAACCAACCAATAGAA 3101
DB 1981 GTTATCTTTTGAATGTTGGTGGGCAACCTCTCTTTGGGCAACCAACCAATAGAA 2040
QY 3102 ACACAAATGAAGTTATCAACTGGCAACATCTCTTCATCTCCACCAAGCAATCAATC 3161
DB 2041 ACACAAATGAAGTTATCAACTGGCAACATCTCTTCATCTCCACCAAGCAATCAATC 2100

QY 3162 AGTCTCTGAAGCTTCTGATCTTATTATAAATTTGCCGAGGACCCGAAGATCGCTTAGGC 3221
DB 2101 AGTCTCTGAAGCTTCTGATCTTATTATAAATTTGCCGAGGACCCGAAGATCGCTTAGGC 2160
QY 3222 AAGAAATGGTGTGATGAATAAAGCTCATCTTTTAAAAACAATGATTCCTCCAGT 3281
DB 2161 AAGAAATGGTGTGATGAATAAAGCTCATCTTTTAAAAACAATGATTCCTCCAGT 2220
QY 3282 GACCTTGAGACAGCTGCTTCTCATACATTCCTAAATACACACACCCACACATACATCA 3341
DB 2221 GACCTTGAGACAGCTGCTTCTCATACATTCCTAAATACACACACCCACACATACATCA 2280
QY 3342 AATTTTGATCTCTGTTGATCTCTGATAAATATGAGTGTGATGAAGAGGAAATGTA 3401
DB 2281 AATTTTGATCTCTGTTGATCTCTGATAAATATGAGTGTGATGAAGAGGAAATGTA 2340
QY 3402 AATGACACTCTCAAT-GGATGGTATAAATAAGGAAAGCTCTGACATGCTATCTATGA 3460
DB 2341 AATGACACTCTCAATGGGATGGTATAAATAAGGAAAGCTCTGACATGCTATCTATGA 2400
QY 3461 ATTTACCTTCCGAAGGTTTTTTTGATGACAAATGGCTACCCATA 3502
DB 2401 ATTTACCTTCCGAAGGTTTTTTTGATGACAAATGGCTACCCATA 2442

RESULT 7

AL583963/C

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

AL583963 Human DNA sequence from clone RPI-203A15 on chromosome 6, complete sequence.
AL583963 HTG.
GI:14547269
human.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 80146)
Hammond, S.
Direct Submission
Submitted (23-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, CB10 1SA, UK. E-mail enquiries: humquery@sanger.ac.uk
requests: clonerequest@sanger.ac.uk
On Jun 25, 2001 this sequence version replaced gi:14133045.
During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em, EMBL; SW, SWISSPROT; Tr, TrEMBL; Wp, WORMPEP; Information on the WORMPEP database can be found at
http://www.sanger.ac.uk/Projects/C_elegans/wormpep This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
http://www.sanger.ac.uk/HGP/Chr6
RPI-203A15 is from the library RPI-1 constructed by the group of Pieter de Jong. For further details see
http://www.chori.org/bacpac/home.htm
VECTOR: pcYFAC2
IMPORTANT: This sequence is not the entire insert of clone RPI-203A15 it may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.

The true left end of clone RP1-317N9 is at 78147 in this sequence.
The true right end of clone RP11-703H16 is at 2000 in this sequence.

FEATURES

source	Location/Qualifiers
repeat_region	14256..15963 /note="L1MA4A repeat: matches 3603..5331 of consensus"
repeat_region	15966..16070 /note="FLAM_C repeat: matches 1..105 of consensus"
repeat_region	16076..16112 /note="Alu repeat: matches 79..160 of consensus"
repeat_region	16122..17029 /note="L1MA4A repeat: matches 5319..6300 of consensus"
repeat_region	17080..17261 /note="MER74A repeat: matches 1..171 of consensus"
repeat_region	17490..17788 /note="AluSg repeat: matches 1..296 of consensus"
repeat_region	17966..18263 /note="AluY repeat: matches 1..298 of consensus"
repeat_region	18502..18551 /note="25 copies 2 mer ta 100% conserved"
repeat_region	18552..18818 /note="AluSg repeat: matches 34..303 of consensus"
repeat_region	18929..19117 /note="MIR repeat: matches 41..241 of consensus"
repeat_region	19252..19557 /note="AluSg repeat: matches 1..306 of consensus"
repeat_region	19558..19843 /note="AluJo repeat: matches 12..294 of consensus"
repeat_region	20971..21270 /note="AluSx repeat: matches 1..300 of consensus"
repeat_region	21302..21391 /note="FRAM/FAM repeat: matches 4..72 of consensus"
repeat_region	22015..22173 /note="MERSB repeat: matches 7..176 of consensus"
misc_feature	22419..23403 /note="CpG island"
repeat_region	22531..22618 /evidence=not_experimental
repeat_region	23378..23453 /note="44 copies 2 mer cc 64% conserved"
repeat_region	23686..23896 /note="MERSA repeat: matches 108..185 of consensus"
repeat_region	24365..24495 /note="FLAM_C repeat: matches 2..125 of consensus"
repeat_region	24659..24859 /note="MER3 repeat: matches 2..209 of consensus"
repeat_region	25072..25249 /note="AluSg repeat: matches 118..290 of consensus"
repeat_region	25256..25572 /note="AluSx repeat: matches 1..309 of consensus"
repeat_region	25575..25874 /note="AluSx repeat: matches 1..299 of consensus"
repeat_region	26374..26417 /note="22 copies 2 mer ta 100% conserved"
repeat_region	27253..27549 /note="AluSg repeat: matches 1..298 of consensus"
repeat_region	27551..27631 /note="LIME repeat: matches 5699..5778 of consensus"
repeat_region	27852..27915 /note="L2 repeat: matches 2203..2267 of consensus"
repeat_region	28020..28320 /note="AluSg repeat: matches 1..305 of consensus"
repeat_region	28891..29184 /note="AluSg repeat: matches 1..292 of consensus"
repeat_region	29260..29548 /note="AluSg repeat: matches 1..305 of consensus"
repeat_region	29567..29652 /note="AluJo/FRAM repeat: matches 202..287 of consensus"
repeat_region	29818..30056 /note="AluSg repeat: matches 1..293 of consensus"
repeat_region	30267..30570 /note="AluSg repeat: matches 1..302 of consensus"
repeat_region	30581..30762 /note="LIME repeat: matches 2353..2196 of consensus"
repeat_region	30827..31119 /note="AluJb repeat: matches 1..312 of consensus"
repeat_region	3174..3476 /note="AluY repeat: matches 2..303 of consensus"
repeat_region	3512..3956 /note="L2 repeat: matches 2264..2750 of consensus"
repeat_region	3975..4127 /note="AluSg/x repeat: matches 145..297 of consensus"
repeat_region	4152..4469 /note="AluSx repeat: matches 1..304 of consensus"
repeat_region	4470..4856 /note="HURS-P3 repeat: matches 6797..7185 of consensus"
repeat_region	4850..6348 /note="HURS-P3 repeat: matches 5120..6625 of consensus"
repeat_region	6349..6647 /note="AluSg repeat: matches 1..301 of consensus"
repeat_region	6648..6815 /note="HURS-P3 repeat: matches 4954..5120 of consensus"
repeat_region	6840..6911 /note="HURS-P3 repeat: matches 4691..4762 of consensus"
repeat_region	6912..6953 /note="Lipba repeat: matches 742..701 of consensus"
repeat_region	6974..7347 /note="HURS-P3 repeat: matches 4568..4942 of consensus"
repeat_region	7349..7384 /note="LIP repeat: matches 1790..1825 of consensus"
repeat_region	7382..7745 /note="Lipba repeat: matches 1590..1958 of consensus"
repeat_region	7745..8418 /note="LIM2 repeat: matches 84..638 of consensus"
repeat_region	8458..8921 /note="LIM2 repeat: matches 677..198 of consensus"
repeat_region	9470..9751 /note="AluSx repeat: matches 12..293 of consensus"
repeat_region	9941..10043 /note="AluJb repeat: matches 210..306 of consensus"
repeat_region	10059..10260 /note="AluSg/x repeat: matches 108..309 of consensus"
repeat_region	10295..10994 /note="LIPB1 repeat: matches 4004..4703 of consensus"
repeat_region	10995..11289 /note="AluSg repeat: matches 1..301 of consensus"
repeat_region	11290..12343 /note="LIPB1 repeat: matches 4703..5780 of consensus"
repeat_region	12344..12479 /note="AluSx repeat: matches 1..142 of consensus"
repeat_region	12480..12769 /note="AluY repeat: matches 1..290 of consensus"
repeat_region	12770..12957 /note="AluSg repeat: matches 142..309 of consensus"
repeat_region	12958..13331 /note="LIPB1 repeat: matches 5779..6152 of consensus"
repeat_region	13382..13962 /note="L1MA4A repeat: matches 3027..3603 of consensus"
repeat_region	13963..14255 /note="AluSc repeat: matches 1..306 of consensus"

QY	725	AGGGAATGTGCAGCAATCAGTTAAACCGCAACAGAGCTGGAAAGTGTCTTAAGAATCCTT	784
Db	57333	AGGGAATGTGCAGCAATCAGTTAAACCGCAACAGAGCTGGAAAGTGTCTTAAGAATCCTT	57274
QY	785	AGTTCCCTCAGAGGATGGCCGCCACCTAGGAGAAGTGTGGCCTATCATCTGAGAGTCC	844
Db	57273	AGTTCCCTCAGAGGATGGCCGCCACCTAGGAGAAGTGTGGCCTATCATCTGAGAGTCC	57214
QY	845	CAACTCAGACAGATGTAGGAAGACCTTTGCTGTGGATCTGGTATATACAGCATTTGTTC	904
Db	57213	CAACTCAGACAGATGTAGGAAGACCTTTGCTGTGGATCTGGTATATACAGCATTTGTTC	57154
QY	905	AGCTCACCTTAGCAACGGACAGAGCTGAACCCGCCACACCACTCAAGTAAGGAGTGT	964
Db	57153	AGCTCACCTTAGCAACGGACAGAGCTGAACCCGCCACACCACTCAAGTAAGGAGTGT	57094
QY	965	TACTCCTCCACCACTCCAGAGGCCAGACTCCCCCTCCAGAGAGTACAACCTCCACCTCC	1024
Db	57093	TACTCCTCCACCACTCCAGAGGCCAGACTCCCCCTCCAGAGAGTACAACCTCCACCTCC	57034
QY	1025	CCCTTCATGGGAACCAACTCTCAACAAAGCGCTATTCTGGAACATGGAAATACGTAAT	1084
Db	57033	CCCTTCATGGGAACCAACTCTCAACAAAGCGCTATTCTGGAACATGGAAATACGTAAT	56974
QY	1085	CTCCCGAATCTCTCTGTCCCACTGGGGCATGCAAGAGGGCTATCCTCCACCACTCT	1144
Db	56973	CTCCCGAATCTCTCTGTCCCACTGGGGCATGCAAGAGGGCTATCCTCCACCACTCT	56914
QY	1145	CAACACTTCCCCCATGAATCCTCTCTAATCAAGGACAGAGGCGATTAGTTCTGTCTCTGT	1204
Db	56913	CAACACTTCCCCCATGAATCCTCTCTAATCAAGGACAGAGGCGATTAGTTCTGTCTCTGT	56854
QY	1205	TGGCAGACACCAATCATATGACAGAGTCTTAGCAAAATTTAACTTCCATCAGGGAGACC	1264
Db	56853	TGGCAGACACCAATCATATGACAGAGTCTTAGCAAAATTTAACTTCCATCAGGGAGACC	56794
QY	1265	TGGAATCGAATGGTACTTGGACAACCTGATTTTCATGATACACCAAAATGTGTCCCTGC	1324
Db	56793	TGGAATCGAATGGTACTTGGACAACCTGATTTTCATGATACACCAAAATGTGTCCCTGC	56734
QY	1325	TGGCAGCTGTAATCGGGCAGCCACCACTCCATATCTCTTGACAGCAGTAATGACAAAG	1384
Db	56733	TGGCAGCTGTAATCGGGCAGCCACCACTCCATATCTCTTGACAGCAGTAATGACAAAG	56674
QY	1385	CCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTTCGTCATATACAAATGGAAGTAT	1444
Db	56673	CCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTTCGTCATATACAAATGGAAGTAT	56614
QY	1445	TCCTCAGCTATGATGTGGCCAAACAGAAATAGTCATTAACATGGAACATATAACATTAG	1504
Db	56613	TCCTCAGCTATGATGTGGCCAAACAGAAATAGTCATTAACATGGAACATATAACATTAG	56554
QY	1505	TGTACCTGGACTGCAAAACAAATGGCCCTCAGTCATCTTCTGCTCCAGCCCACTCATCCCC	1564
Db	56553	TGTACCTGGACTGCAAAACAAATGGCCCTCAGTCATCTTCTGCTCCAGCCCACTCATCCCC	56494
QY	1565	GAGCAGTGGGCATGAAATCCCTACATGGCAACCTAAACATACCACTGAGGTCAAAATCTTT	1624
Db	56493	GAGCAGTGGGCATGAAATCCCTACATGGCAACCTAAACATACCACTGAGGTCAAAATCTTT	56434

PR PI TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, YASUYUKI NISHIYAMA, PI HIDEYUKI SAYA, PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC G01N33/532/ C12P21/02, PC C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00.

(C12N15/00, PC C12R1:91)	
CC Strandedness: Double;	
FH Key	
FT source	
FEATURES	
source	Location/Qualifiers
	1..1374
	/organism="unidentified"
BASE COUNT	442 a 252 c 288 g 392 t
ORIGIN	
Query Match 34.18; Score 1360.4; DB 6; Length 1374;	
Best Local Similarity 99.98; Pred. No. 6.3e-290;	
Matches 1372; Conservative 0; Mismatches 1; Indels 1; Gaps 1;	
Qy	2130 TATTCTCTCAAGCATTTAAATCTTTATGAGCAACATGTAGAAAATGTACTCAAAATCT 2189
Db	1 TATTCTCTCAAGCATTTAAATCTTTATGAGCAACATGTAGAAAATGTACTCAAAATCT 60
Qy	2190 CATCAGAGCGCTACATCGTAAACAAATAGAAATGAATGATCGGGTTGGATTA 2249
Db	61 CATCAGAGCGCTACATCGTAAACAAATAGAAATGAATGATCGGGTTGGATTA 120
Qy	2250 TCTCAAGATGCCAGGATCAATGAGAAAGATGCTTCCCAAAAGAAATCTAAATACATC 2309
Db	121 TCTCAAGATGCCAGGATCAATGAGAAAGATGCTTCCCAAAAGAAATCTAAATACATC 180
Qy	2310 CGTCTTAAAGGGCTTAAATGACAAAGTCTATGTTTGAAGATAAAGACACTAGGAATA 2369
Db	181 CGTCTTAAAGGGCTTAAATGACAAAGTCTATGTTTGAAGATAAAGACACTAGGAATA 240
Qy	2370 GGAGCATTTGGTGAAGTCTGTACGAGAAAGTAGACTAAGCGTTTGTATGCAACA 2429
Db	241 GGAGCATTTGGTGAAGTCTGTACGAGAAAGTAGACTAAGCGTTTGTATGCAACA 300
Qy	2430 AAAACTCTTCCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGSCTGAG 2489
Db	301 AAAACTCTTCCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGSCTGAG 360
Qy	2490 AGAGATATCTGGCTGAGAGCTGACAAATGAATGGGTAGTTCGTCTATATATTCATCCAA 2549
Db	361 AGAGATATCTGGCTGAGAGCTGACAAATGAATGGGTAGTTCGTCTATATATTCATCCAA 420
Qy	2550 GATAAGGCAATTTATCTTTGTAATGACTACATCTCTGGGGGTGATATGATGAGCCTA 2609
Db	421 GATAAGGCAATTTATCTTTGTAATGACTACATCTCTGGGGGTGATATGATGAGCCTA 480
Qy	2610 TTAATTAGAATGGGCATCTTCCGAAAGCTGCGACAGATTTACATAGCAGAACTTACC 2669
Db	481 TTAATTAGAATGGGCATCTTCCGAAAGCTGCGACAGATTTACATAGCAGAACTTACC 540
Qy	2670 TGTGAGTTGAAAGTTCATATAAATGGTTTATTATAGAGATATTAACCTGATAT 2729
Db	541 TGTGAGTTGAAAGTTCATATAAATGGTTTATTATAGAGATATTAACCTGATAT 600
Qy	2730 ATTTTGAATGATCGTGGTGCATATTAATTAAGTACTGACTTTGGCCTCTGCACTGGCTC 2789
Db	601 ATTTTGAATGATCGTGGTGCATATTAATTAAGTACTGACTTTGGCCTCTGCACTGGCTC 660
Qy	2790 AGATGACACAGATTTCTAAGTACTATCAGAGTGGTGACCATCCACGGCAAGATAGCATG 2849
Db	661 AGATGACACAGATTTCTAAGTACTATCAGAGTGGTGACCATCCACGGCAAGATAGCATG 720
Qy	2850 GATTTCAGTAATGAATGGGGGATCCCTCAAGCTGTCGATGTGGAGACAGACTGAAGCCA 2909
Db	721 GATTTCAGTAATGAATGGGGGATCCCTCAAGCTGTCGATGTGGAGACAGACTGAAGCCA 780
Qy	2910 TTAGAGCGGAGAGCTGCACGCCAGCAGCAGCATGCTTAGCACATCTTTGGTGGGACT 2969
Db	781 TTAGAGCGGAGAGCTGCACGCCAGCAGCAGCATGCTTAGCACATCTTTGGTGGGACT 840

RESULT 9	
AF119846/c	
ID	AF119846 standard; RNA; HUM; 1807 BP.
XX	AF119846;
AC	AF119846;
XX	AF119846.1
SV	AF119846.1
XX	12-MAY-2000 (Rel. 63, Created)
DT	09-MAY-2001 (Rel. 67, Last updated, Version 2)
DE	Homo sapiens PRO1474 mRNA, complete cds.
XX	
KW	
OS	Homo sapiens (human)
OC	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia;
OC	Eutheria; Primates; Catarrhini; Hominidae; Homo.
XX	[1]
RP	1-1807
RA	Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RT	"Functional prediction of the coding sequences of 79 new genes deduced by
RL	analysis of cDNA clones from human fetal liver";
XX	Unpublished.
XX	[2]
RP	1-1807
RA	Zhang C., Yu Y., Zhang S., Wei H., Zhang Y., Zhou G., Bi J., Liu M., He F.;
RT	Submitted (13-JAN-1999) to the EMBL/GenBank/DBJ databases.
RL	Department of Experimental Hematology, Institute of Radiation Medicine,
RL	Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
XX	SPTREMBL; Q9P199; Q9P199.
DR	
XX	

Key	Location/Qualifiers
FH	1. .1807
FT	/db_xref="taxon:9606"
FT	/organism="Homo sapiens"
FT	/clone="FLB5626"
FT	/tissue_type="liver"
FT	/dev_stage="fetus"
FT	complement(841..1530)
FT	/codon_start=1
FT	/db_xref="SPTREMBL:Q9P199"
FT	/note="predicted protein of H01474"
FT	/product="PRO1474"
FT	/protein_id="AAF69600.1"
FT	/translation="MADKLSRVKASYTFQEVHVFIOGLPEGLMAGLQVPEVONLS
FT	DWPFELMLVHALAQSEVCVROVGEGLQODLGGDLGKLGKGVGLVOLQDSQVGFQIIG
FT	IFGLRLYIALQGOVLRIVPVDLEQLAHVGSLLKHIRHDFCLPQPLAKWRRRR
FT	WPGVGRARSURLGRAALANGSPVRLPALPHOGYLLGLLRTLAASASALRPVAAVGE
FT	LCSPLOS"
XX	
SQ	Sequence 1807 BP; 452 A; 440 C; 451 G; 464 T; 0 other;
Query Match	23.4%; Score 930.8; DB 17; Length 1807;
Best Local Similarity	99.8%; Pred. No. 5.6e-195;
Matches 932; Conservative	0; Mismatches 2; Indels 0; Gaps 0;
QY	3 CTTTGGTTGCTGGGAGGACTCTGGCGCCTCAGGCTCCGCCCTCAGCGCCGTGGCGGC 62
DB	
QY	934 CTTGGGTTGCTGGGAGGACTCTGGCGCCTCAGGCTCCGCCCTCAGCGCCGTGGCGGC 875
DB	
QY	63 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTATTTATTTATTTAAAC 122
DB	
QY	874 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTATTTATTTATTTAAAC 815
DB	
QY	123 AGTCTGGGAGCTCTTGAAGGATCATTTTCACATTTTCTCAGAGAAAGCTCTGGATC 182
DB	
QY	814 AGTCTGGGAGCTCTTGAAGGATCATTTTCACATTTTCTCAGAGAAAGCTCTGGATC 755
DB	
QY	183 TATCAAAATAAGAAAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGACT 242
DB	
QY	754 TATCAAAATAAGAAAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGACT 695
DB	
QY	243 GAAAGCCAGAGGATATAGACAAATAGAGGCTTAAGACCTTTCTGCCAGTAACTATCT 302
DB	
QY	694 GAAAGCCAGAGGATATAGACAAATAGAGGCTTAAGACCTTTCTGCCAGTAACTATCT 635
DB	
QY	303 GTCAGTAGCCGCAATGTTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCA 362
DB	
QY	634 GTCAGTAGCCGCAATGTTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCA 575
DB	
QY	363 TCTGATGCTGCTAAGGCTGAGCATACATGAGTAAATGCTCAACCGAAGATCCTCGACAA 422
DB	
QY	574 TCTGATGCTGCTAAGGCTGAGCATACATGAGTAAATGCTCAACCGAAGATCCTCGACAA 515
DB	
QY	423 GTCAGAAATCCACCAAAATTTGGGAGCGCATATAAGGCTTTCAGAGAAATTCGAACCTCT 482
DB	
QY	514 GTCAGAAATCCACCAAAATTTGGGAGCGCATATAAGGCTTTCAGAGAAATTCGAACCTCT 455
DB	
QY	483 CTGCTTCCATTTGCAAAATGAACAAATTTCTTCGAGTACTTTCAGAGTAAATTCACAA 542
DB	
QY	454 CTGCTTCCATTTGCAAAATGAACAAATTTCTTCGAGTACTTTCAGAGTAAATTCACAA 395
DB	
QY	543 ATGCTTCAAGACTCTCAAGCTCTGGATTTGATGAGGATATGTTTATCAAGCTCTTCAG 602
DB	
QY	394 ATGCTTCAAGACTCTCAAGCTCTGGATTTGATGAGGATATGTTTATCAAGCTCTTCAG 335
DB	
QY	603 AAACTAACACAGAGTATAGAACGACCAATTTGAATTCATTTAGTAAATGAGTTTACCAA 662
DB	
QY	334 AAACTAACACAGAGTATAGAACGACCAATTTGAATTCATTTAGTAAATGAGTTTACCAA 275
DB	
QY	663 GATCTTCGACGAGACAGATGCTGCGAGCTGCCAGACCTATTAAATGCCAGATGAA 722
DB	
QY	274 GATCTTCGACGAGACAGATGCTGCGAGCTGCCAGACCTATTAAATGCCAGATGAA 215
DB	

QY	723	CCAGGAATGTGCAGCAATCAGTTAAACCGCAACAGAGCTGGAAGGTTCTTAAAGATCC	782
DB	214	CCAGGAATGTGCAGCAATCAGTTAAACCGCAACAGAGCTGGAAGGTTCTTAAAGATCC	155
QY	783	TTAGTTCCTCAGAGGATGCGCCGCTAGGAGAAAGTGTGGCCTATCATTTCTGAGAGT	842
DB	154	TTAGTTCCTCAGAGGATGCGCCGCTAGGAGAAAGTGTGGCCTATCATTTCTGAGAGT	95
QY	843	CCCAACTCACACAGATGTAGGAAGACCTTTGCTGGATCTGGTATATATCATCATTTGTT	902
DB	94	CCCAACTCACACAGATGTAGGAAGACCTTTGCTGGATCTGGTATATATCATCATTTGTT	35
QY	903	CAAGCTCACCTTAGCAACGACGACAGAGTGAACC	936
DB	34	CAAGCTCACCTTAGCAACGACGACAGAGTGAACC	1
RESULT 10			
AC102712/c			
LOCUS	AC102712	164490 bp	DNA linear HTG 21-AUG-2002
DEFINITION	Mus musculus clone RP24-258P4, WORKING DRAFT SEQUENCE, 17 unordered pieces.		
ACCESSION	AC102712		
VERSION	AC102712.2	GI:22381704	
KEYWORDS	HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	1 (bases 1 to 164490)		
AUTHORS	Birren, B., Nusbaum, C. and Lander, E.		
TITLE	Mus musculus, clone RP24-258P4		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 164490)		
AUTHORS	Birren, B., Linton, L., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Boguslavsky, L., Boukhalter, B., Brown, A., Camarata, J., Campiano, A., Chang, J., Chazaro, B., Choepel, Y., Colangelo, M., Collins, S., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzHugh, W., Gage, D., Galagan, J., Gardyna, S., Ginde, S., Gord, S., Goyette, M., Graham, L., Grand-Pierre, N., Hages, B., Heaford, A., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., LaRocque, K., Lanazares, R., Landers, T., Lehoczy, J., Levine, R., Liu, G., Maclean, C., Macdonald, P., Major, J., Marquis, N., Matthews, C., McCarthy, M., McEwan, P., McKernan, K., McPheeters, R., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K., Phunkhang, P., Pierre, N., Pollara, V., Raymond, C., Retta, R., Rieback, M., Riley, R., Rise, C., Rogov, P., Roman, J., Rosetti, M., Roy, A., Santos, R., Schauer, S., Schupback, R., Seaman, S., Severy, P., Spencer, B., Stange-Thomann, N., Stojanovic, N., Straus, N., Subramanian, A., Talamas, J., Tesfaye, S., Theodore, J., Topham, K., Travers, N., Travis, N., Trigilio, J., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Ye, W. J., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.		
TITLE	Direct Submission		
JOURNAL	Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA		
REFERENCE	3 (bases 1 to 164490)		
AUTHORS	Birren, B., Nusbaum, C., Lander, E., Ali, A., Allen, N., Anderson, S., Barna, N., Bastien, V., Bloom, T., Boguslavsky, L., Boukhalter, B., Camarata, J., Chang, J., Chazaro, B., Choepel, Y., Collamore, A., Cook, A., Cooke, P., Dearellano, K., Dewar, K., Diaz, J. S., Dodge, S., Faro, S., Ferreira, P., FitzGerald, M., Gage, D., Galagan, J., Gardyna, S., Gord, S., Graham, L., Grand-Pierre, N., Hages, B., Horton, L., Hulme, W., Iliev, I., Johnson, R., Jones, C., Kanat, A., Karatas, A., Kells, C., Landers, T., Levine, R., Lindblad-Toh, K., Liu, G., Maclean, C., Macdonald, P., Major, J., Matthews, C., McCarthy, M., Meldrim, J., Meneus, L., Mihova, T., Mlenga, V., Murphy, T., Naylor, J., Nguyen, C., Nicol, R., Norbu, C., Norman, C. H., O'Connor, T., O'Donnell, P., O'Neil, D., Oliver, J., Peterson, K.,		

Phunkhang, P., Pierre, N., Raymond, C., Retta, R., Rise, C., Rogov, P., Roman, J., Roy, A., Schauer, S., Schuback, R., Seaman, S., Severy, P., Smith, C., Spencer, B., Stange-Thomann, N., Stojanovic, N., Talamas, J., Testaye, S., Theodore, J., Topham, K., Travers, M., Vassiliev, H., Viel, R., Vo, A., Wilson, B., Wu, X., Wyman, D., Young, G., Zainoun, J., Zembek, L., Zimmer, A. and Zody, M.

TITLE JOURNAL

Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA

On Aug 21, 2002 this sequence version replaced gi:17060822.

All repeats were identified using RepeatMasker:

Smit, A.F.A. & Green, P. (1996-1997)

http://ftp.genome.washington.edu/RM/RepeatMasker.html

Genome Center

Center: Whitehead Institute/ MIT Center for Genome Research

Center code: WIBR

Web site: http://www.seq.wi.mit.edu

Contact: sequence_submissions@genome.wi.mit.edu

Project Information

Center project name: L19333

Center clone name: 258_P4

Summary Statistics

Sequencing vector: Plasmid; n/a; 100% of reads

Chemistry: Dye-terminator Big Dye; 100% of reads

Assembly program: Phrap; version 0.960731

Consensus quality: 158929 bases at least Q40

Consensus quality: 161241 bases at least Q30

Consensus quality: 162174 bases at least Q20

Insert size: 176000; agarose-fp

Quality coverage: 7.2 in Q20 bases; agarose-fp

Quality coverage: 7.8 in Q20 bases; sum-of-contigs

NOTE: This is a 'working draft' sequence. It currently

consists of 17 contigs. The true order of the pieces

is not known and their order in this sequence record is

arbitrary. Gaps between the contigs are represented as

runs of N, but the exact sizes of the gaps are unknown.

This record will be updated with the finished sequence

as soon as it is available and the accession number will

be preserved.

1 1303: contig of 1303 bp in length

1304 1403: gap of 100 bp

1404 2163: contig of 760 bp in length

2164 2263: gap of 100 bp

2264 3531: contig of 1268 bp in length

3532 3631: gap of 100 bp

3632 4716: contig of 1085 bp in length

4717 4816: gap of 100 bp

4817 8090: contig of 3274 bp in length

8091 8190: gap of 100 bp

8191 11116: contig of 2926 bp in length

11117 11216: gap of 100 bp

11217 41036: contig of 29820 bp in length

41037 41136: gap of 100 bp

41137 45062: contig of 3926 bp in length

45063 45162: gap of 100 bp

45163 48629: contig of 3467 bp in length

48630 48729: gap of 100 bp

48730 57775: contig of 9046 bp in length

57776 57875: gap of 100 bp

57876 70851: contig of 12976 bp in length

70852 70951: gap of 100 bp

70952 85997: contig of 15046 bp in length

85998 86097: gap of 100 bp

86098 104897: contig of 18800 bp in length

104898 104997: gap of 100 bp

104998 120197: contig of 15200 bp in length

120198 120297: gap of 100 bp

120298 136805: contig of 16508 bp in length

136806 136905: gap of 100 bp

136906 160104: contig of 23199 bp in length

160105 160204: gap of 100 bp

* 160205 164490: contig of 4286 bp in length.

FEATURES Source

Location/Qualifiers

1..164490

/organism="Mus musculus"

/db_xref="taxon:10090"

/clone="RP24-258P4"

/clone_lib="RPCI-24 Male Mouse BAC"

1..1303

/note="assembly_fragment"

1404..2163

/note="assembly_fragment"

2264..3531

/note="assembly_fragment"

3632..4716

/note="assembly_fragment"

4817..8090

/note="assembly_fragment"

8191..11116

/note="assembly_fragment"

11217..41036

/note="assembly_fragment"

41137..45062

/note="assembly_fragment"

45163..48629

/note="assembly_fragment"

48730..57775

/note="assembly_fragment"

57876..70851

/note="assembly_fragment"

70952..85997

/note="assembly_fragment"

86098..104897

/note="assembly_fragment"

104998..120197

/note="assembly_fragment"

120298..136805

/note="assembly_fragment"

136906..160104

/note="assembly_fragment"

160205..164490

/note="assembly_fragment"

clone_end:T7

vector_side:right"

BASE COUNT 47330 a 35286 c 34742 g 45505 t 1627 others

ORIGIN

Query Match 21.3%; Score 848.4; DB 2; Length 164490;

Best Local Similarity 86.8%; Pred. No. 1e-176;

Matches 957; Conservative 0; Mismatches 141; Indels 4; Gaps 2;

QY 725 AGGGAATGTCAGCAATCAGTTAAACCGCAACAGAGCTGAAAGGTTCTAAAGAACCTT 784

Db 5920 AGGAATGTCAACATTCAATTAACCGCAACAGCTGAAAGGTTCTAAAGAGTCTCT 5861

QY 785 AGTTCCTCAGAGGATGCGCCGCCACTAGGAGAAAGTGGCCTATCATCTCAGAGTCC 844

Db 5860 AGTTCCTCAGAGACACGCGCCACTCTCTAGGAGAAATGTGTTATCGTTCTGAAGCCC 5801

QY 845 CAACTCAGACAGATCTAGGAGAACCTTTGCTGGATCTGGTATATCAGCATTTGTTC 904

Db 5800 CAACTCAGCGGATGTAGGAGACCTCTCTGGATCGGCATTCGACATTTGCTCA 5741

QY 905 AGCTCACCTAGCAACGAGAGAGAGTGAACCCGCCACACACCTCAAGTAAGAGTGT 964

Db 5740 AGCTCACCAAGCAATGGACAGAGAGTGAACCCGCCACACCTCAAGTAAGAGTGT 5681

QY 965 TACTCCTCCACCTCCAGAGCGCCAGCTCCCTCCAGAGAGTCAACTCCACCTCC 1024

Db 5680 TACTCCTCCACCTCCAGAGCGCCAGCTCCCTCCAGAGAGTCAACTCCACCTCC 5621

QY 1025 CCCTTTCATGGGAACCAAACTCTCAAAACAAAGCGCTATTCTGGAACATGGAATACGTA 1084

Db 5620 CCCTTTCATGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAATACGTA 5561

PD	06-APR-1999
PF	24-SEP-1997 JP 1997258689
PR	
PI	TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
YASUYUKI NISHIYAMA,	
PI	HIDEYUKI SAYA
PC	C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC
G01N33/53Z/C12P21/02,	
PC	(C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,
(C12N15/00, PC C12R1:91)	
CC	Strandedness: Doublet;
CC	Topology: Linear;
FH	Key
FT	source
FT	Location/Qualifiers
FT	1..795 /organism='Unidentified'
FEATURES	Location/Qualifiers
source	1..795
	/organism="unidentified"
	/db_xref="taxon:32644"
BASE COUNT	224 a 228 c 160 g 183 t
ORIGIN	
	Query Match 19.8%; Score 788.6; DB 6; Length 795;
	Best Local Similarity 99.5%; Pred. No. 1.5e-163;
	Matches 791; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY	786 GTTCTCTAGAGGGATGCGCGGCACCTAGGAGAAGTGTCGCCATTATCTTGAGAGTCCC 845
DB	1 GTTCTCTAGAGGGATGCGCGGCACCTAGGAGAAGTGTCGCCATTATCTTGAGAGTCCC 60
QY	846 AACTCACACACATGTAGGAAGACCTTTCTCTGGATCTGGTATATCAGCATTTGTTCAA 905
DB	61 AACTCACACACATGTAGGAAGACCTTTCTCTGGATCTGGTATATCAGCATTTGTTCAA 120
QY	906 GCTCACTTAGCAACGGACAGAGTGAACCCCCACCACCTCAAGTAAGAGTGT 965
DB	121 GCTCACTTAGCAACGGACAGAGTGAACCCCCACCACCTCAAGTAAGAGTGT 180
QY	966 ACTCTTCCACCACTCCAAGAGGCCAGACTTCCCCCTCCAAGAGGTACAACCTCCACCTCC 1025
DB	181 ACTCTTCCACCACTCCAAGAGGCCAGACTTCCCCCTCCAAGAGGTACAACCTCCACCTCC 240
QY	1026 CCTTCATGGGAACCAACTCTCAAAACAAGCGCTATCTGGAAAACATGGNAATCGTAATC 1085
DB	241 CCTTCATGGGAACCAACTCTCAAAACAAGCGCTATCTGGAAAACATGGNAATCGTAATC 300
QY	1086 TCCCGAATCTCTCTGTCCACCTGGGGATGCAAGAGSGCTATCTCCACCACCTCTC 1145
DB	301 TCCCGAATCTCTCTGTCCACCTGGGGATGCAAGAGSGCTATCTCCACCACCTCTC 360
QY	1146 AACACTTCCCCCAATGAATCCTCTTAATCAAGGACAGAGAGGCATTAGTTCTGTCTCTGTT 1205

1206	QY	GGCAGACAACCAATCATCATGCGAGATTCTAGCAAAATTTAACTTTCCATCAGGGAGACCT	1265
421	Db	GGCAGACAACCAATCATCATGCGAGATTCTAGCAAAATTTAACTTTCCATCAGGGAGACCT	480
1266	QY	GGAAATCGAAGTGGTACTGGACAAACTGATTTTCATGATACACAAAAATGTTGTCCTGCCT	1325
481	Db	GGAAATCGAAGTGGTACTGGACAAACTGATTTTCATGATACACAAAAATGTTGTCCTGCCT	540
1326	QY	GGCACTGTGAATCGGACGAGCCACCACTCCATATCTCTGACAGCAGCTAATGGACAAAGC	1385
541	Db	GGCACTGTGAATCGGACGAGCCACCACTCCATATCTCTGACAGCAGCTAATGGACAAAGC	600
1386	QY	CCTTCCTGTTTACAAACAGGGGATCTGCTGCTCCTTCGTCTATATACAAATGGAAGTATT	1445
601	Db	CCTTCCTGTTTACAAACAGGGGATCTGCTGCTCCTTCGTCTATATACAAATGGAAGTATT	660
1446	QY	CCTCAGCTCATGATGGTGGCAACAGAAATAGTCATAAAGATGGAACATATATAACATTAGT	1505
661	Db	CCTCAGCTCATGATGGTGGCAACAGAAATAGTCATAAAGATGGAACATATATAACATTAGT	720

```
QY 1506 GTACCTGGAGTCAAAACAATGGCGCTCAGTCATCTTCTGCTCCAGCCAGTCATCCCG 1565
DB 721 GTACCTGGAGTCAAAACAATGGCGCTCAGTCATCTTCTGCTCCAGCCAGTCATCCCG 780
QY 1566 AGCAGTGGGCATGAA 1580
DB 781 AGCAGTGGGCATGAA 795

RESULT 12
LOCUS AR201457 3155 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 7 from patent US 6359193.
ACCESSION AR201457
VERSION AR201457.1 GI:20252345
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
  1 (bases 1 to 3155)
  Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
  Nucleotide sequences of lats genes
  Patent: US 6359193-A 7 19-MAR-2002.
FEATURES
  source
    1..3155
    /organism="unknown"
BASE COUNT 751 a 924 c 894 g 586 t
ORIGIN
  Query Match 19.1%; Score 760; DB 6; Length 3155;
  Best Local Similarity 63.7%; Pred. No. 3.2e-157;
  Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

QY 1553 CACTCTGCTAATCTCAGCCCTTCTGCTACACAGTCACTGCAATTCACACAGCTCCTATT 1712
DB 985 CAACCTGAGCCCTCACTGCGCCGCCCAACACAGCTCACCCGCTGAGCGCCGACACATC 1044
QY 1713 CAACAGCCTGTGAAAGTATGCTGTATTAAACACAGAGTACAGACTGCTTTAGCACCT 1772
DB 1045 CTTACCCCTGTGAAGACGTGCTGCTGCGGCCGAGCCGACAGCAGCGTGGGCC 1104
QY 1773 ACACACCTTTCTTGGATACACAGCCAAATTCAAACCTGTTCAACCCAGTCTCTTTCTGAG 1832
DB 1105 TCGCACCCCGCTTGGTGGCTGCGCCCAACAGCCTGCCACTGAGAGCTGGAGAGCAAG 1164
QY 1833 GGAACCCCTTCAATGTGACTGTGATGCCACCTGTTCTGAAGTCCAAACTATCAAGGA 1892
DB 1165 GAGGCGAGCGCAGCCGCCACACCCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGGTGC 1224
QY 1893 CCACACACCCCTACCCAAACATCTGCTGCACCAACCCATCTGTTCTCTCCATACGAG 1952
DB 1225 CCACCGCTCGTATCCAAAGCACTTCTGCTGCCAGTAACTGAGCAGTACAGCGTG 1284
QY 1953 TCAATCAGTAAAGCTAGCAAGAGGATACGCCAAGCTTGCCCAAGGAAGATGAGAGTGAA 2012
DB 1285 GACCTGGACAGCCCTGTCACCAAGTGTGACAGAGTCTGCGAGGGGCACTGTACTAGAC 1344
QY 2013 AAGAGTTA-----TGAATGTTCATAGTGGGATGAAGAAAGAAACAG 2057
DB 1345 GGGAGTGACAGAGCCCAAGAGTGGCAAGGGAGGAGCAAAAGCTGGCAGAGCAAAAGCAG 1404
QY 2058 ATTAACAATCTCACCTATTACTTGTAGGAAACAAAGAAAGATGAAGAGCGAAGGAATCT 2117
DB 1405 ATTACAGCTCCCGGTGCTGTCGCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
QY 2118 CGTATTCAAAGTTATCTCCTCAAGCAATTTAAATCTTTTATGGAGCAACATGTAGAAAT 2177
DB 1465 CGCATCAAGAGTACTCCCTTATGCCCTTCAATTTCTCATGGAGCAACAGCTGGAGAAAT 1524
QY 2178 GTACTCAATCTCATCAGCAGCTCATACGTGTAAGAAACAAATTAGAAGATGAATATG 2237
DB 1525 GTCATCAAAACCTTACCAGCAAGGTGAGCGGAGGTACAGCTGGAGCAGGAATGGCC 1584
```

```
QY 2238 CGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAGAA 2297
DB 1585 AAAGCTGGGCTCTGTGAGCCGAGCAGGAGAGATGAGGAAGATCTCTTACGAGAAGGAG 1644
QY 2298 TCTAATTACATCCGCTTTAAAGGGCTAAATGGACAGTCTATGTTGTGAAGATAAG 2357
DB 1645 TCTAATACACCGGCTGAAGAGGCCAAGATGGCAAGTCCATGTTGTGAAAATCAAG 1704
QY 2358 ACACCTAGGATAGGAGCATTTGTTGAAGTCTGTCTAGCAAGAAAGATAGATCTAAGGCT 2417
DB 1705 ACTTAGGATCGGTGCTTTGGGAAGTGTGCTGCTGTTGAAGTGGACACTCACGCT 1764
QY 2418 TTGTATGCAACAAAACCTTCTGAAAGAAAGATGTTCTTCTCGAAATCAAGTGCCTAT 2477
DB 1765 CTGTAGCCCATGAAGACTCTCAGGAAGAGGATGCTCTCAACCGGAATCAAGTGGCCAT 1824
QY 2478 GTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACAAATGAATGGTAGTCTGCTATAT 2537
DB 1825 GTCAGGCTGAGAGGAGCATCTGCTGAAGCAGCAATGAGTGGGTGCTAAACTCTAC 1884
QY 2538 TATTCATTCCAAGATAAGGACAATTTATCTTGTAAATGGACTACATTTCTTGGGGTAT 2597
DB 1885 TACTCTTCCAGCAAGGACAGCCTGTACTTGTGATGGACTACATACAGCGGGGAT 1944
QY 2598 ATGATGAGCCTATTAAATAGAATGGCATCTTTCCAGAAAGTCTGGCAGATTTACATA 2657
DB 1945 ATGATGAGCCTGCTGATCAGGATGAGGCTCTTCCCTGAGCACCTGGCCGCTTCTCAT 2004
QY 2658 GCAGAACTTACCTGCTGAGTGAAGTGTTCATAAATGGTGTATTTATTCATAGAGATAT 2717
DB 2005 GCAGAGTTGACCTGGCCATTTGAAGTGTCCAAAGTGGCTTTATCCACCGGACATC 2064
QY 2718 AAACCTGATAAATTTTGTATGATCGTGTGATGTCATATTAATTTGACTGACTTTGGCCTC 2777
DB 2065 AAGCTTGCAACAATCTATCTGACCTGGATGGTGTGATTAAGCTGACAGATTTGGCCCTC 2124
QY 2778 TGCACTGGCTTTCAGATGGACACAGATTTCTAAGTACTATCAGAGTGTGACCATCCACGG 2837
DB 2125 TGCACTGGATTTCAAGTGGAGTCAATTTCAAGTACTACCAAGAGGGAACCAATGAGA 2184
QY 2838 CAAGATAGCATGATTTCAAGTGAATGAATGGGGGATCCCTCAAGCTGTCGATCTGGAGAC 2897
DB 2185 CAGGACAGATGAGGCGCGGTGACCTCTGGGAGATGTTTCCAACTGTCGCTGTGGAGAC 2244
QY 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCAGCAGCAGAGTGTCTAGCAATCT 2957
DB 2245 AGCTTAAAGACCTTGAGCAGAGAGGGCGCAGACGACGAGGTGCTTGGCAGATCT 2304
QY 2958 TTGTTGGGACTCCCAATTTATTTGACCTGAAAGTGTGCTAGCAAGATACACACAG 3017
DB 2305 CTTGTGGGACACAAATTTACATCGCTCCGAGGTGCTTCTCCGCAAGGGTACACGCGAG 2364
QY 3018 TCTGTGATTTGGAGTGTGGTGTATTCTTTTGAATGTTGGTGGGACAACTCTCT 3077
DB 2365 CTCTGTGACTGGTGGAGCGTGGTGTGATTTCTTTTGTAGATGCTGTTGGGAGCGCCCT 2424
QY 3078 TTCTTGGCACAACACCAATTTAGAAACACAAATGAAGTTATCAACTGGGCAACATCTCTT 3137
DB 2425 TTCTTGGCCCCCACCACACAGAGCAGCTCAAGGTGATCAACTGGGAGAGCAGCTG 2484
QY 3138 CACATTCACCAACAGCTAACTCAGTCTCCTGAGCTTCTGATCTTATTATTAACTTTCG 3197
DB 2485 CATATCCCTACGAGGTGAGGCTCAGCGTGGGCGCCGAGACCTCATCAGAGAGCTGTC 2544
QY 3198 CGAGGACCCCAAGATCGCTTTAGCAAGAATGGTCTGCTGATAAATAAAGCTCATCAATTT 3257
DB 2545 TGGCGGCTGACTGCCCTGGGCGGATGGGCGAGATGACCTCAAGGACACACCGCTC 2604
QY 3258 TTTAAAACAAATTCCTCAGTACAGCTGAGAGCAGAGTCTGCTTCATACATTCCTTAA 3317
DB 2605 TTCAACACCATCGACTTTTCCCGTGACATCCGAAAGCAGGCTGACACCTACGCTCCACC 2664
```


RESULT 14			
E38226			PAT 31-JAN-2002
LOCUS	E38226	5486 bp	linear
DEFINITION	Human tumor regulatory gene;		

<hr/>					
ACCESSION	E38226				
VERSION	E38226.1	GI:	18626935		
KEYWORDS	JP 2000210086-A/1.				
SOURCE	Homo sapiens				
ORGANISM	Homo sapiens				
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Theria; Primates; Catarrhini; Hominidae; Homo.					
REFERENCE	1 (bases 1 to 5486)				
AUTHORS	Koga,J., Kono,K. and N,Z.F.				
TITLE	Human tumor regulatory gene				
JOURNAL	Patent: JP 2000210086-A 1 02-AUG-2000;				
	NITON CHEMICAL RESEARCH K K				
COMMENT	OS Homo sapiens (human) PN JP 2000210086-A/1 PD 02-AUG-2000 PF 25-JAN-1999 JP 1999016223 PR				
	PI JUNICHI KOGA, KEIKO KONO, ZOROTARYOFU FIODORU N PC C12N15/09,A61K31/00,A61K38/27,C07K14/47,C07K14/525,C12Q1/68,PC G01N33/50, PC G01N33/50,G01N33/574,G01N33/577//C12P21/08,(C12N15/09,PC C12R1:91),C12N15/00, CC A61K37/36,(C12N15/00,C12R1:91) CT				
	Key	Location/Qualifiers			
PH	source	1..5486			
FT		/organism='Homo sapiens (human)'			
FEATURES	source				
	1..5486				
	/organism="Homo sapiens"				
	/db_xref="taxon:9606"				
BASE COUNT	1378 a 1452 c 1390 g 1266 t				
ORIGIN					
Query Match 18.7%; Score 744.4; DB 6; Length 5486;					
Best Local Similarity 66.3%; Pred. No. 9.1e-154;					
Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps					
Qy	1872	GAAGTCCTCAAACTATCAAGGACCACCAACCCCTTACCCAAAACATCTGCT-----GCACC	1926		
Dd	1905	GGAGGCCACCGGAGGTGCGCCGCTCCGCGCTACCCGAAGCACCTGCTGCTGCGCAGC	1964		
Qy	1927	AAAAACCATCTGTTCTCCATACAGACTCAATCAGTAGCTAGCCCTAGCAAGAGGATCAG-	1982		
Dd	1965	AAGTCGGAGCAGTACGACCTGGACGCGCTGTGCCAGGCGATGGAGCAGCCCTCCGTGCG	2024		
Qy	1983	--CCAGCTTGCCCCAAGAAGATGAGAGTCAAAGAGTTATGAAAATTGTGTAGTAGGG	2039		
Dd	2025	GCCCCAACGAGCCGAGGCGCGCACNAGAGCCGAACGCCCAAGGGGCAAAAGGC	2084		
Qy	2040	GATAAGAAAAAGAACAGATTACAACCTTCACTTACTGTGTAGGAAAAACAAGAAAGAT	2099		
Dd	2085	GGAAGGATAAAAAGCAGATTTCAGACCTCTCCGCTCCCGCCCAAAAACAGCAGAGAC	2144		
Qy	2100	GAAGACCAAGGGAATCTCGTATTCAAGTATTCTCCTCAAGCATTTAAATTTCTTTATG	2159		
Dd	2145	GAAGAGAGAGAGAGTACGCGATCAAGAGCTACTCGCCNTAGCCTTTAAGTCTTCTATG	2204		
Qy	2160	GAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGCGCTTACATCGTAAACAA	2219		
Dd	2205	GAGCAGCAGCTGGAGAATGTCAATAAACCTACCAGCAGAAGGTTAACCGGAGCTGCAG	2264		
Qy	2220	TTAGAGATGAATGATCGGGTTGGATTATCTCAAGATGCCAGGATCAATCAGAAAG	2279		
Dd	2265	CTGGAGCAAGAAATGCCCAAGCTGGACTCTGTGAAGCTGAGCAGGACGATCGGAG	2324		
Qy	2280	ATGCTTTTGCRAAAAGATCTAAATTACATCCGCTCTTAAAGGGGTAAATGGCAAGTCT	2339		
Dd	2325	ATTCCTTACCAGAAAGTCTTAATTACACAGGTTAAGAGGCGCAGATGCACAAGTCT	2384		
Qy	2340	ATGTTTGTGAAGATAAAGACACTAGGAATAGAGCATTTTGTGTAAGTCTGTCTAGCAGA	2399		
Dd	2385	ATGTTTGTCAAGATCAAAACCTGGGGATCGGTGCTTTTGGAGAGTGTGCGTCTGTGT	2444		

Db 2205 GAGCAGACGCTGGAGATGTCTATCAAAACCTACCAGCAGAGGTTTAAACCGAGGCTGCAG 2264
 Qy 2220 TTACAGAAATGAATGATGCGGGTTGGATTATCTCAAGATCCCGAGGATCAAAATGAGAAAG 2279
 Db 2265 CTGGAGCAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATGCGGAAG 2324
 Qy 2280 ATGCTTTGCCAAAAGAAATCTAATTAGATCCGCTTTAAAGGGCTTAAATGGCAAGTCT 2339
 Db 2325 ATCTCTTACCAGAAAGAGTCTAATTACAAAGGTTAAAGAGGGCCAAAGATGGACAGTCT 2384
 Qy 2340 ATGTTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTCTTAGCAAGA 2399
 Db 2385 ATGTTTGTCAAGATCAAAACCTGGGGATCGGTGCTTTGGAGAGTGTGCCITGCTTGT 2444
 Qy 2400 AAAGTAGATACTAAGGCTTTGTATGCAACAAAACCTCTTGAAGAAAGATGTCTTCTT 2459
 Db 2445 AAGGTGGACACTACGCCCTGTACGCCATGAAGACCCCTAAGGAAAAGAGATGCTCTGAAC 2504
 Qy 2460 CGAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCTGGCTGAAGCTGCAATGAA 2519
 Db 2505 CGGAATCAGGTGGGCCACGCTCAAGCCGAGAGGACATCTGGCCGAGGACAGCAATGAG 2564
 Qy 2520 TGGTAGTTCCTATATATATTCATTCACAGATAAGGACAAATTTATATCTTTGTAATGAC 2579
 Db 2565 TGGTGGTCAAACTCTACTACTCTTCCAAGCAAAAGACGCTGTACTTTTGTGATGGAC 2624
 Qy 2580 TACATTCTGGGGTGATATGATGAGCCTTAAATAGAAATGGGCATCTTCCAGAAAGT 2639
 Db 2625 TACATCCCTGGTGGGACATGATGAGCCTGCTGATCCGGATGGAGGCTTCCCTGAGCAC 2684
 Qy 2640 CTGGCACCATTCTACATAGCAGAACTTACCTGTGCAGTTGAAGTGTTCATATAAATGGT 2699
 Db 2685 CTGCCCCGGTCTACATCGCAGAGCTGACTTTGGCCATTGAGAGTGTCCCAAGATGGC 2744
 Qy 2700 TTTATTCATAGAGATTAACCTGATTAATTTTGTATGATGATGATGATGATGATTAATA 2759
 Db 2745 TTTATTCACCGAGACATCAAGCCTGATTAACATTTTATGATAGATCTGGATGGTCAATATA 2804
 Qy 2760 TTGACTGACTTTGGCTCTGACCTGGCTTCAGATGGACACAGATTTCAAGTACTATCAG 2819
 Db 2805 CTCACAGATTTGGGCTCTGCACTGGGTTCAGGTGGATGCAATTCACAAATATACCAG 2864
 Qy 2820 AGTGTGTACCATCCACGGCAAGATAGCATTTTCATTAATGAATGGGGGATCCCTCA 2879
 Db 2865 AAAGGACCCATGTGACAGACAGCAGATGGAGCCAGCCTCTGGGATGATGTGTCT 2924
 Qy 2880 AGCTGTGATGAGAGACAGACTGAGCCATAGAGCGGAGCTGACAGCCAGCAGCAGCAG 2939
 Db 2925 AACTGTGCTGTGGGACAGGCTGAAGACCCCTAGAGAGGGGCGGAGAGGAGTGTCTCTC 2984
 Qy 2940 CGATGCTAGCACATCTTTGGTGGGACTCCCAATTTATTTGACCTGAAGTGTGTCTA 2999
 Db 2985 AGTGTGCTGGACATTTCACTGGTGGGACTCCAACTACATCGACCCGAGGAGTGTCTCTC 3044
 Qy 3000 CGAACAGGATACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTTTTCGAATG 3059
 Db 3045 CGCAAAAGGTACACTCAACTCTGTGACTGGTGGAGTGTGGAGTGTATTTCTTCGAGATG 3104
 Qy 3060 TTGTTGGGACACCTCTCTTCTTGGCAAAACACCATTAAGAACACAAATGAAGTTATC 3119
 Db 3105 CTGGTGGGGGACGCCCTTTTGGCACTCTCCACAGAACCCAGCTGAAGTGTATC 3164
 Qy 3120 AACTGGCAAACTCTTTTACATTTCCACCAAGTAACTCAGTCTTGAAGCTTCTGAT 3179
 Db 3165 AACTGGGAGAACGCTCCACATTTCCAGCCAGGTGAAGCTGAGCCCTGAGGCCAGGGAC 3224
 Qy 3180 CTTATTATTAACCTTTGGGAGGACCCGAGATCGCTTAGGCAAGAAATGGTCTGATGAA 3239
 Db 3225 CTCATCACCAAGCTGTGTGCTCCGACAGACCACCCCTGGGGCGAATGGGGCCGATGAC 3284
 Qy 3240 ATAAAGCTCATCCATTTTAAACAAATTTGACTTCTCCAGTGAAGTGAAGAGAGTCT 3299
 Db 3285 CTGAGGCCACCCCTCTTTCAGGCCATTTGACTCTCCAGTGACATCCGGAAGCAGCCA 3344

Qy 3300 GCTTTCATACATTCCTAAAATCACACACCCCAAGATACATCAAAATTTTGTCTGTGAT 3359
 Db 3345 GCCCCCTAGTTCCCAACCATCAGCCACCCCATGGACACCTCGAATTTCCAGCCCGTAGAT 3404
 Qy 3360 CTTGATAAAATTTATGGAGTGTATGATAACGAGGAGAAAATGTAAATGACACTCTCAATGGA 3419
 Db 3405 GAAGAAAGCCCTTGGAAACGATGCCAGC---GAAGGTAGCACCACCAAGGCTTGGGACACACTC 3461
 Qy 3420 TGGTATAAAATGGAAAGCATCCTGAACATGCATTTCTATGAATTTTACCTTCCGAAGGTTT 3479
 Db 3462 ACCTCGCCCAATAAACAGCATCCTGAGCAGCATTTTACGAATTTACCTTCCGAAGGTTT 3521
 Qy 3480 TTTGATGACAATGGCTACCCATTAATTTCCGAAGCCTATTGAATATGAATACATTAAT 3539
 Db 3522 TTTGATGACAATGGCTACCCCTTCGATGCCCAAGGCTTCAGGACGAGAAGCTTCACAG 3581
 Qy 3540 TCACAAGGCTCAGA 3553
 Db 3582 GCTGAGAGCTCAGA 3595

Search completed: January 16, 2003, 18:10:29
 Job time : 7470.05 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:07 ; Search time 508.163 Seconds
(without alignments)
17655.667 Million cell updates/sec

Title: US-09-763-334-1
Perfect score: 3984
Sequence: 1 accttgggtgctgggacg.....agagttttgaaatttttt 3984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 2185239 seqs, 1125999159 residues
Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :	N_Geneseq_101002:*
1:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT:*
2:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT:*
3:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT:*
4:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT:*
5:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT:*
6:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT:*
7:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT:*
8:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT:*
9:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT:*
10:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT:*
11:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT:*
12:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT:*
13:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT:*
14:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT:*
15:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT:*
16:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT:*
17:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT:*
18:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT:*
19:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT:*
20:	/SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT:*
21:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT:*
22:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT:*
23:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT:*
24:	/SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3984	100.0	3984	17	AAT42118
2	3984	100.0	3984	21	AAZ51505
3	3978.8	99.9	3982	20	AAZ87396
4	2483	62.3	3213	20	AAZ51506
5	2479.8	62.2	3213	17	AAT42119
6	2422	60.8	2442	20	AAZ32981
7	1360.4	34.1	1374	20	AAZ32982
8	1233.2	31.0	1357	22	ABA08740
9	788.6	19.8	795	20	AAZ32984

10	760	19.1	3155	17	AAT42120	M-lats2 gene encod
11	760	19.1	3155	21	AAZ51507	Mouse Lats2 (large
12	744.4	18.7	5276	20	AAZ87397	Human WART2 CDNA.
13	744.4	18.7	5486	21	AAZ51506	DNA encoding a tum
14	744.4	18.7	5486	21	AAZ51506	DNA encoding a tum
15	742.8	18.6	1961	21	AAZ51506	cDNA SS1771A encod
16	741.8	18.6	3533	22	AAI71765	Human cancer-inhib
17	693.8	17.4	1912	23	ABK43457	DNA encoding novel
18	675.8	17.0	1498	21	AAZ61158	cDNA SS1771 encodi
19	655.4	16.5	2043	22	AAZ03989	Human protein tyro
20	609.8	15.3	638	21	AAZ79942	Human colon cancer
21	580.4	14.6	582	20	AAZ32983	Human warts gene f
22	478.6	12.0	1501	22	AAZ75341	Human TGF-beta rec
23	469	11.8	5720	17	AAT42117	Lats gene encoding
24	469	11.8	5720	21	AAZ51508	Drosophila melanog
25	467.4	11.7	3319	23	ABL03169	Drosophila melanog
26	452.8	11.4	678	21	AAZ79982	Human colon cancer
27	346.2	8.7	676	22	AAZ27179	cDNA encoding nove
28	346.2	8.7	676	23	ABK43776	DNA encoding novel
29	269.4	6.8	1935	20	AAZ06834	Disease associated
30	269.4	6.8	4983	22	AAZ44629	Novel protein kina
31	254.6	6.4	3018	17	AAT31452	Human Ndr serine/t
32	254.6	6.4	3018	24	ABK84773	Human cDNA differe
33	254.6	6.4	3800	22	ABA08668	Human NDR homologu
34	233	5.8	1710	21	AAZ43403	Arabidopsis thalia
35	233	5.8	2003	21	AAZ49841	Arabidopsis thalia
36	218.2	5.5	2001	21	AAZ36174	Arabidopsis thalia
37	216.8	5.4	11187	23	ABL03168	Drosophila melanog
38	210.6	5.3	1894	21	AAZ39567	Arabidopsis thalia
39	209	5.2	1922	21	AAZ49426	Arabidopsis thalia
40	186.4	4.7	734	23	ABK43709	DNA encoding novel
41	186	4.7	734	22	AAZ27161	cDNA encoding nove
42	186	4.7	734	22	AAZ56728	Human immune/haema
43	186	4.7	734	23	ABK43985	DNA encoding novel
44	180.4	4.5	2255	23	ABL29695	Drosophila melanog
45	173.2	4.3	2160	22	AAI66703	A. gossypii AG007

ALIGNMENTS

RESULT 1	
AAZ42118	
ID	AAT42118 standard; cDNA; 3984 BP.
AC	AAT42118;
XX	
DT	31-JAN-1997 (first entry)
DE	H-lats gene encoding large tumour suppressor.
XX	
KW	Human; h-lats gene; large tumour suppressor; fetal brain;
KW	protein-serine/threonine-kinase; cell proliferation; antisense;
KW	dominant-negative; cancer; degenerative disorder; trauma;
KW	growth deficiency; therapy; antitumour; vulnary; diagnostic;
KW	transgenic plant; transgenic animal; growth; senescence; ds.
XX	
OS	Homo sapiens.
XX	
Key	Location/Qualifiers
FT	231..3623
FT	/*tag= a
FT	/product= Lats protein
XX	
XX	WO6G30402-A1.
XX	
PD	03-OCT-1996.
XX	
PF	26-MAR-1996; 96WO-US04101.
XX	
PR	27-MAR-1995; 95US-0411111.
XX	
PA	(UYVA) UNIV YALE.

XX Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX WPI; 1996-455275/45.
 DR P-PSDB; AAW05178.
 XX
 XX New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX
 XX Claim 6; Page 118-123; 215pp; English.
 XX
 CC This sequence encodes a human large tumour suppressor h-lats protein,
 CC and has been isolated from a fetal human brain phage lambda-gt10
 CC cDNA library using a 2.1-kb DNA probe from the mouse m-lats gene
 CC (AAT42119). The full-length cDNA is present in plasmid pBS(KS)-h-lats.
 CC The gene encodes a putative protein-serine/threonine-kinase, and
 CC inhibits cell proliferation and plays a crucial role throughout
 CC development. Activators or inhibitors of lats function (e.g. an
 CC antisense oligonucleotide or dominant-negative lats fragment) may be
 CC used in therapy of cancer or other proliferative disorders,
 CC degenerative disorders, trauma, growth deficiency, etc., and
 CC fragments of the gene may be used as diagnostic probes. A
 CC lats-inhibitor sequence may be expressed in a transgenic plant or
 CC farm animal to confer increased growth and inhibit senescence.
 XX
 SQ Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;
 Query Match 100.0%; Score 3984; DB 17; Length 3984;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 ACCTTTGGTGTGGGACGACTCTGGCGCCCTCAGCGTCCGCCCTCAGCGCCGCTGGCC 60
 DB 1 ACCTTTGGTGTGGGACGACTCTGGCGCCCTCAGCGTCCGCCCTCAGCGCCGCTGGCC 60
 QY 61 GCTCTCAGGAGCTCTGCTCTCCCTCCAGAGTTAATTTATTTATTTAAGAAATTTTA 120
 DB 61 GCTCTCAGGAGCTCTGCTCTCCCTCCAGAGTTAATTTATTTATTTAAGAAATTTTA 120
 QY 121 ACAGTCTCTGGGACTCTCTCAAGGATCATTTTCACATTTTCTCAGAGAAAGCTCTGGA 180
 DB 121 ACAGTCTCTGGGACTCTCTCAAGGATCATTTTCACATTTTCTCAGAGAAAGCTCTGGA 180
 QY 181 TCTATCAATAAAGAGTCTCTGTGGGTACATATATAGATGTTTTCATGAGAGGA 240
 DB 181 TCTATCAATAAAGAGTCTCTGTGGGTACATATATAGATGTTTTCATGAGAGGA 240
 QY 241 GTGAAAGCCAGAGGATATAGCAATGAGGCTTAAGACCTTTCTCCAGTAACATA 300
 DB 241 GTGAAAGCCAGAGGATATAGCAATGAGGCTTAAGACCTTTCTCCAGTAACATA 300
 QY 301 CTGTCACTAGCGGCAATGTTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360
 DB 301 CTGTCACTAGCGGCAATGTTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360
 QY 361 CATCTGATGCTGCTAAGGCTGAGCATACATGAGTAAATGTAACCGAAGATCCTCGAC 420
 DB 361 CATCTGATGCTGCTAAGGCTGAGCATACATGAGTAAATGTAACCGAAGATCCTCGAC 420
 QY 421 AAGTCAGAAATCCACCAATTTGGGCGCATCATAAAGCCCTTCAGGAAATTCGAAC 480
 DB 421 AAGTCAGAAATCCACCAATTTGGGCGCATCATAAAGCCCTTCAGGAAATTCGAAC 480
 QY 481 CTCTGCTTCCATTTGCAATGAAACAATTTCTCTCGAGTACTTCAGAGTAAATCCAC 540
 DB 481 CTCTGCTTCCATTTGCAATGAAACAATTTCTCTCGAGTACTTCAGAGTAAATCCAC 540
 QY 541 AAATGCTTCAAGACTTCAAGCTGCTGGATTTGATGAGGATATGTTATACAGCTCTTC 600
 DB 541 AAATGCTTCAAGACTTCAAGCTGCTGGATTTGATGAGGATATGTTATACAGCTCTTC 600
 QY 601 AGAAACTTAACACAGAGTATAGACGCAATTTGAATTCATTAGTAAATGAGTTACC 660
 DB 601 AGAAACTTAACACAGAGTATAGACGCAATTTGAATTCATTAGTAAATGAGTTACC 660

DB 601 AGAAACTTAACACAGAGTATAGACGCAATTTGAATTCATTAGTAAATGAGTTACC 660
 QY 661 AAGATCTCTCAGCAGAGCAGATGGCTGCAGAGCTGCCAGACCTATTAAATGCCAGCATGA 720
 DB 661 AAGATCTCTCAGCAGAGCAGATGGCTGCAGAGCTGCCAGACCTATTAAATGCCAGCATGA 720
 QY 721 AACAGGGAATGTGCAGCAATCAGTTAACCCGCAACAGAGCTGGAAAGGTTCTAAAGAT 780
 DB 721 AACAGGGAATGTGCAGCAATCAGTTAACCCGCAACAGAGCTGGAAAGGTTCTAAAGAT 780
 QY 781 CCTTAGTCTCTCAGAGCAGTGGCCGCCCTAGGAGAAAGTGTGCCCTATCATTTCTGAGA 840
 DB 781 CCTTAGTCTCTCAGAGCAGTGGCCGCCCTAGGAGAAAGTGTGCCCTATCATTTCTGAGA 840
 QY 841 GTCCCAACTCACAGACAGATGTAGGAACACCTTTCTGTGGATCTGTATATCAGCATTTG 900
 DB 841 GTCCCAACTCACAGACAGATGTAGGAACACCTTTCTGTGGATCTGTATATCAGCATTTG 900
 QY 901 TTCAAGCTCACCCCTAGCAACGAGAGTGAACCCGCCACCCACCTCAAGTAAGGA 960
 DB 901 TTCAAGCTCACCCCTAGCAACGAGAGTGAACCCGCCACCCACCTCAAGTAAGGA 960
 QY 961 GTGTACTCTCCACCACCTCCAGAGGCCAGACTCCGCCCTCCAGAGGTACAATCCAC 1020
 DB 961 GTGTACTCTCCACCACCTCCAGAGGCCAGACTCCGCCCTCCAGAGGTACAATCCAC 1020
 QY 1021 CTCCCCCTTCATGGGAACCAAACTCTCAACAAAGCGCTATTCTGGAACATGGAATACG 1080
 DB 1021 CTCCCCCTTCATGGGAACCAAACTCTCAACAAAGCGCTATTCTGGAACATGGAATACG 1080
 QY 1081 TAATCTCCGAAATCTCTCTCCACCTGGGGGACGAGAGGGCTATCTCTCCACCAC 1140
 DB 1081 TAATCTCCGAAATCTCTCTCCACCTGGGGGACGAGAGGGCTATCTCTCCACCAC 1140
 QY 1141 CTCTCAACACTTCCCCCATGAATCTCTCTCAATCAAGGACAGAGAGGCTATTAGTTCTGTC 1200
 DB 1141 CTCTCAACACTTCCCCCATGAATCTCTCTCAATCAAGGACAGAGAGGCTATTAGTTCTGTC 1200
 QY 1201 CTGTTGGCAGACAACCAATCATGACAGTCTTAGCAAAATTTAACTTTCCATCAGGGA 1260
 DB 1201 CTGTTGGCAGACAACCAATCATGACAGTCTTAGCAAAATTTAACTTTCCATCAGGGA 1260
 QY 1261 GACCTGGAATCGAAGTGTGACAACTGATTTTCATGATACACCAAAATTTGTTGCC 1320
 DB 1261 GACCTGGAATCGAAGTGTGACAACTGATTTTCATGATACACCAAAATTTGTTGCC 1320
 QY 1321 CTGCTGGCAGTGAATCGGAGCCACCCTCCATATCTCTGACAGCAGCTAATGAC 1380
 DB 1321 CTGCTGGCAGTGAATCGGAGCCACCCTCCATATCTCTGACAGCAGCTAATGAC 1380
 QY 1381 AAAGCCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTCTGTCATATACAAATGGA 1440
 DB 1381 AAAGCCCTTCTGCTTTACAAACAGGGGATCTGCTGCTCTCTGTCATATACAAATGGA 1440
 QY 1441 GTATTCTCAGTCTATGATGTGCAACAGAAATAGTACATAACATGGAACATATATACA 1500
 DB 1441 GTATTCTCAGTCTATGATGTGCAACAGAAATAGTACATAACATGGAACATATATACA 1500
 QY 1501 TTAGTGTACCTGGACTGCAAACTTGGCTCAGTCACTTTCTCTCCAGCCAGTCAAT 1560
 DB 1501 TTAGTGTACCTGGACTGCAAACTTGGCTCAGTCACTTTCTCTCCAGCCAGTCAAT 1560
 QY 1561 CCCGAGCAGTGGGATGAAATCCCTACATGGCAACCTTAACATCCAGTGGGTCAAAT 1620
 DB 1561 CCCGAGCAGTGGGATGAAATCCCTACATGGCAACCTTAACATCCAGTGGGTCAAAT 1620
 QY 1621 CTTTAAATACCCATTAGGAATAGCAAGTCACTCTGCTTAATTTCTCAGGCTTCTGCTA 1680
 DB 1621 CTTTAAATACCCATTAGGAATAGCAAGTCACTCTGCTTAATTTCTCAGGCTTCTGCTA 1680
 QY 1681 CAACAGTCACTGCAATACACAGCTCTCTATTCAACAGCCCTGTGAAAAGTATGCTGTAT 1740
 DB 1681 CAACAGTCACTGCAATACACAGCTCTCTATTCAACAGCCCTGTGAAAAGTATGCTGTAT 1740

QY	1741	TAAACACAGAGCTACAGACTGCTTTAGCACCTTACACACCCCTTCTTGGATACCACAGCAA	1800
DB	1741		1800
QY	1801	TAAACACAGAGCTACAGACTGCTTTAGCACCTTACACACCCCTTCTTGGATACCACAGCAA	1860
DB	1801		1860
QY	1801	TTCAAACTGTTCAACCCAGTCCTTTTCTGAGGGAAACCGCTTCAAACTGTGACTGTGATGC	1860
DB	1801		1860
QY	1861	CACCTGTTGCTGAAGCTCCAAACTATCAAGGACCACCCACCCCTACCCAAAACATCTGC	1920
DB	1861		1920
QY	1921	TGCACAAAACCCATCTGTTCTCCATACGAGTCAATCAGTAAAGCTAGCAAGAGGATC	1980
DB	1921		1980
QY	1981	AGCCAAGCTTGCCCAAGGAAGTGAAGTGAAGAGGTTATGAAAATGTTGATAGTGGG	2040
DB	1981		2040
QY	2041	ATAAGAAAAGAACACAGATTACAACCTTCACTTACTGTTAGGAAAACAAAGAAAGATG	2100
DB	2041		2100
QY	2101	AAGAGCAAGGGAACTCTGTAATTCAAAGTTATCTCTCAAGCATTTAAATCTTTATGG	2160
DB	2101		2160
QY	2161	AGCAACATGTAGAAAATGTACTCAAAATCPCATCAGCAGCTCTACATCGTAAAAACAAT	2220
DB	2161		2220
QY	2221	TAGAAATGAAATGTATGCGGGTTGGATTATCTCAAGATGCCCGAGATCAAAATGAGAAAGA	2280
DB	2221		2280
QY	2281	TGCTTTGCCAAAAGAAATCTAATTACATCCGCTTTAAAGGGCTTAAATGGACAAGTCTA	2340
DB	2281		2340
QY	2341	TGTTTGTGAAGATTAAGACACTAGGAATAGGAGCATTTGGTGAAGCTCTGTCTAGCAAGAA	2400
DB	2341		2400
QY	2401	AAGTAGATCTAAGGCTTTGTATGCAACAAAACCTCTTCGAAAGAAAGATGTTCTTCTTC	2460
DB	2401		2460
QY	2461	GAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTGCGTGAAGCTGACAAATCAAT	2520
DB	2461		2520
QY	2521	GGGTAGTCTCTATATTATTCATCCAAAGATAAGGACAAATTTATCTTTGTAATGGACT	2580
DB	2521		2580
QY	2581	ACATTCTGGGGTGATATGATCAGGCTATTAATTAGAAATGGGCATCTTTCCAGAAAGTC	2640
DB	2581		2640
QY	2641	TGGCAGCATCTACATAGCAGAACTTACCTGTGCAAGTTGAAAGTGTTCATAAAATGGGTT	2700
DB	2641		2700
QY	2701	TTATTCTAGAGATATTAACCTGTAAATTTTGTGATTCGTGATGGTGCATATTAAT	2760
DB	2701		2760
QY	2761	TGACTGACTTTGGCCTCTGCATGGCTTCAGATGACACACAGATCTAAAGTACTATCAGA	2820
DB	2761		2820

QY	2821	GTGTGACCATCCACGCCAAGATAGCATGGATTTTCAGTAATGAATGGGGGATCCCTCAA	2880
DB	2821	GTGTGACCATCCACGCCAAGATAGCATGGATTTTCAGTAATGAATGGGGGATCCCTCAA	2880
QY	2881	GCTGTCGATGTGGAGACAGACTCAAGCCATTAGACGGAGAGCTGCACGCCAGCACCCAGC	2940
DB	2881	GCTGTCGATGTGGAGACAGACTCAAGCCATTAGACGGAGAGCTGCACGCCAGCACCCAGC	2940
QY	2941	GATGCTTAGCACATCTCTTTGGTTGGGACTCCCAATATATATGACCTGAAGTGTGCTAC	3000
DB	2941	GATGCTTAGCACATCTCTTTGGTTGGGACTCCCAATATATATGACCTGAAGTGTGCTAC	3000
QY	3001	GAACAGGATACACAGTGTGTGTGATGTGGAGTGTGGTGTATCTCTTTTGAATGT	3060
DB	3001	GAACAGGATACACAGTGTGTGTGATGTGGAGTGTGGTGTATCTCTTTTGAATGT	3060
QY	3061	TGGTGGGACAACTCTCTTTGGGACAAACACCACTATAGAAACACAAATGAAGTTATCA	3120
DB	3061	TGGTGGGACAACTCTCTTTGGGACAAACACCACTATAGAAACACAAATGAAGTTATCA	3120
QY	3121	ACTGGCAAAACATCTCTTACATCTCCACCAAGCTAAACTCAGTCTCGAAGCTTCTGATC	3180
DB	3121	ACTGGCAAAACATCTCTTACATCTCCACCAAGCTAAACTCAGTCTCGAAGCTTCTGATC	3180
QY	3181	TATATTATAAAGTTGCCGAGGACCCGAAGATCGCTTAGGCAAGAATGGTCTGATGAAA	3240
DB	3181	TATATTATAAAGTTGCCGAGGACCCGAAGATCGCTTAGGCAAGAATGGTCTGATGAAA	3240
QY	3241	TAAAGCTCATCTATTTTTTAAACAAATGACTTCTCCAGTGACCTGAGACAGCAGTCTG	3300
DB	3241	TAAAGCTCATCTATTTTTTAAACAAATGACTTCTCCAGTGACCTGAGACAGCAGTCTG	3300
QY	3301	CTTCATACATCTCTTAAATCACACACCACACAGATACATCAAAATTTTGATCCTGTGATC	3360
DB	3301	CTTCATACATCTCTTAAATCACACACCACACAGATACATCAAAATTTTGATCCTGTGATC	3360
QY	3361	CTGATAAATTTATGGAGTGATGATGACGAGGAGAGAAAATGTAATGACACTCTCAATGGAT	3420
DB	3361	CTGATAAATTTATGGAGTGATGATGACGAGGAGAGAAAATGTAATGACACTCTCAATGGAT	3420
QY	3421	GGTATAAAATGGAAGGACCTCGAACATGCAATCTATGAATTTTACCTCCGAAGGTTTT	3480
DB	3421	GGTATAAAATGGAAGGACCTCGAACATGCAATCTATGAATTTTACCTCCGAAGGTTTT	3480
QY	3481	TTGATGACAATGGCTACCCATATAATTATCCGAAGCCTATTGAATATGAATACATAAAT	3540
DB	3481	TTGATGACAATGGCTACCCATATAATTATCCGAAGCCTATTGAATATGAATACATAAAT	3540
QY	3541	CACAAGGCTCAGACAGCAGTGGAGTGAAGATGATCAAAACACAGGCTCAGAGATTAATA	3600
DB	3541	CACAAGGCTCAGACAGCAGTGGAGTGAAGATGATCAAAACACAGGCTCAGAGATTAATA	3600
QY	3601	ATCCGGATCTAGTATATGTTTACACACTAGTAATAATGTAATGAGGATTTCTGTAAG	3660
DB	3601	ATCCGGATCTAGTATATGTTTACACACTAGTAATAATGTAATGAGGATTTCTGTAAG	3660
QY	3661	GGCTGAAATCGGAGGTGTTTTGAGGTTCTGAGAGTAAAAATATGCAAAATATGACAGC	3720
DB	3661	GGCTGAAATCGGAGGTGTTTTGAGGTTCTGAGAGTAAAAATATGCAAAATATGACAGC	3720
QY	3721	TATATATGTGCTGTGTGACAAATATTTTTTCTCTAAAAATATGGGAAATCCTTTTAA	3780
DB	3721	TATATATGTGCTGTGTGACAAATATTTTTTCTCTAAAAATATGGGAAATCCTTTTAA	3780
QY	3781	AACTGTTAAATTTATCCAGCGGTTTAAATCAGTATTTAGAAAAAAATTTGTTATAGGAAG	3840
DB	3781	AACTGTTAAATTTATCCAGCGGTTTAAATCAGTATTTAGAAAAAAATTTGTTATAGGAAG	3840
QY	3841	TAAATATGAACCTGAATATTTATAGTCAGTCTCTGTGACTTAAAGTACTTAAAAATAAGTAG	3900
DB	3841	TAAATATGAACCTGAATATTTATAGTCAGTCTCTGTGACTTAAAGTACTTAAAAATAAGTAG	3900
QY	3901	TGCTTTCTTTTAAAGGAGAAACCTCGTACTATTTTGTATATATGCTAAATTAATTTTAAAA	3960

Db 3901 TGCCTTTGTTTAAAGGAGAACCTGGTATCTATTGCTATATATGCTAAATAATTTAAAA 3960
 QY 3961 TACAAGAGTTTTTCAAAATTTTTT 3984
 Db 3961 TACAAGAGTTTTTCAAAATTTTTT 3984

RESULT 2

AAZ51505

ID AAZ51505 standard; DNA: 3984 BP.

XX AC AAZ51505;

XX XX

DT 21-JUN-2000 (first entry)

XX Human Lats (large tumour suppressor) DNA.

XX XX

XX Human; Lats; large tumour suppressor; cytostatic; vulnery;

XX cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;

XX treatment; prevention; screening; cancer; skin; ovarian tumour;

XX soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;

XX LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;

XX dysplasia; degenerative disorder; growth deficiency; physical trauma;

XX hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX OS Homo sapiens.

XX XX

FH Key Location/Qualifiers

FT CDS 231..3623

FT /*tag= a

FT /product= "Lats protein"

PN WO200010602-A1.

XX

PD 02-MAR-2000.

XX

PF 18-AUG-1999; 99WO-US19068.

XX

PR 18-AUG-1998; 98US-0096996.

XX

PR 18-AUG-1998; 98US-0096997.

XX

PA (UYVA) UNIV YALE.

XX

PI Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

PI Turenchalk GS, Stewart RA;

XX

DR WPI; 2000-246496/21.

XX

DR P-PSDB: AAY70390.

XX

PT Use of lats proteins, complexes of lats and cdc2 for treating cancer

PT that is refractory to treatment by standard chemotherapy and radiation

PT therapy, and disorders associated with aberrant levels of cdc2 activity

XX

PS Claim 44; Fig 12; 134pp; English.

XX

CC The present sequence is a DNA encoding human Lats (large tumour

CC suppressor) protein which is a cell overproliferation inhibitor and a

CC negative regulator of cell cycle-dependent kinase cdc2/cyclin A.

CC The present sequence is useful for treating cancer that is refractory

CC to standard chemotherapy or radiation therapy such as hyperplasia,

CC metaplasia, or dysplasia, and disorders associated with aberrant

CC levels of cdc2 activity. Conditions treated by promoting cdc2 function

CC include degenerative disorders, growth deficiencies, hypoproliferative

CC disorders, physical trauma, lesions, and wounds. An animal model

CC preferably a mouse, in which a lats gene has been disrupted by homologous

CC recombination, e.g. a lats knock-out mouse, is used for screening

CC compounds that can be used to treat or prevent cancer, particularly

CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders

CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)

XX hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

SQ Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;

Query Match

Best Local Similarity 100.0%; Score 3984; DB 21; Length 3984;

Matches 3984; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ACCTTTGGGTTGCTGGAGCGGACTCTGGCGGCTCAGGCTCGCGCTCAGGCCCGCTGGCC 60
 Db 1 ACCTTTGGGTTGCTGGAGCGGACTCTGGCGGCTCAGGCTCGCGCTCAGGCCCGCTGGCC 60

QY 61 GCTGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATTTATTTAAGAAATTTTA 120
 Db 61 GCTGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATTTAAGAAATTTTA 120

QY 121 ACAGTCTTGGGAGCTTCTTGAAGGATCATTTTTCATTTTCTCAGAAAGAGTCTTGA 180
 Db 121 ACAGTCTTGGGAGCTTCTTGAAGGATCATTTTTCATTTTCTCAGAAAGAGTCTTGA 180

QY 181 TCTATCAATAAAGAGTCTTCTGCTGGGCTACATATATAGATGTTTTCATGAGAGGA 240
 Db 181 TCTATCAATAAAGAGTCTTCTGCTGGGCTACATATATAGATGTTTTCATGAGAGGA 240

QY 241 GTGAAAGCCAGAGGATATACAAATAGAGCCCTTAAGACCTTTCTGCCCAAGTAACTATA 300
 Db 241 GTGAAAGCCAGAGGATATACAAATAGAGCCCTTAAGACCTTTCTGCCCAAGTAACTATA 300

QY 301 CTGTCAAGTACCGCGCAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360
 Db 301 CTGTCAAGTACCGCGCAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC 360

QY 361 CATCTGATGCTGCTAAGGCTGAGCATTAACATGAGTAAATGTCAACCGAAGATCCTCGAC 420
 Db 361 CATCTGATGCTGCTAAGGCTGAGCATTAACATGAGTAAATGTCAACCGAAGATCCTCGAC 420

QY 421 AAGTCAGAAATCCACCAATTTGGGAGCGCATATAAAGCCTTTCAGAGAAATTCGAAACT 480
 Db 421 AAGTCAGAAATCCACCAATTTGGGAGCGCATATAAAGCCTTTCAGAGAAATTCGAAACT 480

QY 481 CTCTGCTTCCATTTGCAAAATCAACAAATTTCTCTCGGAGTACTTCAGAAAGTTAATCCAC 540
 Db 481 CTCTGCTTCCATTTGCAAAATCAACAAATTTCTCTCGGAGTACTTCAGAAAGTTAATCCAC 540

QY 541 AAATGCTTCAAGACTTGCAGAGTCTGCTGATTTGATGAGGATATGCTTATACAAAGCTTTC 600
 Db 541 AAATGCTTCAAGACTTGCAGAGTCTGCTGATTTGATGAGGATATGCTTATACAAAGCTTTC 600

QY 601 AGAAACTAACACAGAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGAGTTACC 660
 Db 601 AGAAACTAACACAGAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGAGTTACC 660

QY 661 AAGATCTTCGACGAGAGCAGATGCTGACAGCTGCGAGACCTTAAATGCCAGCATGA 720
 Db 661 AAGATCTTCGACGAGAGCAGATGCTGACAGCTGCGAGACCTTAAATGCCAGCATGA 720

QY 721 AACGAGGAATGTCAGCAATCAGTTTAAACGCAACAGAGCTGGAAGGTTCTTAAAGAAT 780
 Db 721 AACGAGGAATGTCAGCAATCAGTTTAAACGCAACAGAGCTGGAAGGTTCTTAAAGAAT 780

QY 781 CCTTAGTCTCTAGAGCATGCGCCGCTAGGAGAAAGTGTGGCTATCATTTCTGAGA 840
 Db 781 CCTTAGTCTCTAGAGCATGCGCCGCTAGGAGAAAGTGTGGCTATCATTTCTGAGA 840

QY 841 GTCCCAACTCACAGACAGATGTAGGAAGACCTTTGCTGGATCTGTATATCAGCAATTTG 900
 Db 841 GTCCCAACTCACAGACAGATGTAGGAAGACCTTTGCTGGATCTGTATATCAGCAATTTG 900

QY 901 TTCAAGCTCACCTTAGCAGCAGAGAGTGAACCCGCCACACCTCAAGTAAAGA 960
 Db 901 TTCAAGCTCACCTTAGCAGCAGAGAGTGAACCCGCCACACCTCAAGTAAAGA 960

QY 961 GTGTACTCTCCACCCTTCAAGAGGCGACACTCCCTCCAGAGGTACAACCTCCAC 1020
 Db 961 GTGTACTCTCCACCCTTCAAGAGGCGACACTCCCTCCAGAGGTACAACCTCCAC 1020

Qy	1021	CTCCCCCTTCATGGGAACCAACTCTCAAAACAAAGCGCTATTCTGGAAACATGGAATACG	1080
Db	1021	CTCCCCCTTCATGGGAACCAAACTCTCAAAACAAAGCGCTATTCTGGAAACATGGAATACG	1080
Qy	1081	TAATCTCCCGAAATCTCTCTGTGCCACCTGGGGCATGCCAAGAGGGCTATCTCTCCACCAC	1140
Db	1081	TAATCTCCCGAAATCTCTCTGTGCCACCTGGGGCATGCCAAGAGGGCTATCTCTCCACCAC	1140
Qy	1141	CTCTCAACACTTCCCCCATGAATCTCTTAATCAAGGACAGAGGCGATTAGTCTCTGTC	1200
Db	1141	CTCTCAACACTTCCCCCATGAATCTCTTAATCAAGGACAGAGGCGATTAGTCTCTGTC	1200
Qy	1201	CTGTTGGCAGACAACCAATCATCTGCAGAGTTCTAGCAAAATTTAACTTTCATCAGGGA	1260
Db	1201	CTGTTGGCAGACAACCAATCATCTGCAGAGTTCTAGCAAAATTTAACTTTCATCAGGGA	1260
Qy	1261	GACCTGGAATGCAAGATGTTACTGGACAACATGATTTTCATGATACACCAAAATGTTCTCC	1320
Db	1261	GACCTGGAATGCAAGATGTTACTGGACAACATGATTTTCATGATACACCAAAATGTTCTCC	1320
Qy	1321	CTGCTGGCACTGTGAATCGGCAGCCACCCTCCATATCTCTGCACAGCAGCTAATGTCAC	1380
Db	1321	CTGCTGGCACTGTGAATCGGCAGCCACCCTCCATATCTCTGCACAGCAGCTAATGTCAC	1380
Qy	1381	AAAGCCCTTCTGCTTTTACAAACAGGGGATCTGCTGCTCCTTTCGTCATATACAAATGGAA	1440
Db	1381	AAAGCCCTTCTGCTTTTACAAACAGGGGATCTGCTGCTCCTTTCGTCATATACAAATGGAA	1440
Qy	1441	GTATTCTCTCAGTCTATGATGGTGCCAAACAGAAATAGTCATAACATGGAATATATAACA	1500
Db	1441	GTATTCTCTCAGTCTATGATGGTGCCAAACAGAAATAGTCATAACATGGAATATATAACA	1500
Qy	1501	TTAGTGTACCTGGACTGCAAAACAAATGGCCCTCAGTCACTCTTCTGCTCCAGCCGAGTCAT	1560
Db	1501	TTAGTGTACCTGGACTGCAAAACAAATGGCCCTCAGTCACTCTTCTGCTCCAGCCGAGTCAT	1560
Qy	1561	CCCGAGCAGTGGGATGAATCCCTACATGCCACCTTAAATACCATACCAGTGAAGTCAAAAT	1620
Db	1561	CCCGAGCAGTGGGATGAATCCCTACATGCCACCTTAAATACCATACCAGTGAAGTCAAAAT	1620
Qy	1621	CTTTTAATTAACCCATTAGGAATAGAGCAAGTCACTCTGCTAAATCTCAGCCTTCTGCTA	1680
Db	1621	CTTTTAATTAACCCATTAGGAATAGAGCAAGTCACTCTGCTAAATCTCAGCCTTCTGCTA	1680
Qy	1681	CAACAGTCTGCAATTACACGAGTCTTATTTCAACAGCCTGTGAAAGATAGCGGTGAT	1740
Db	1681	CAACAGTCTGCAATTACACGAGTCTTATTTCAACAGCCTGTGAAAGATAGCGGTGAT	1740
Qy	1741	TAAACCCAGAGCTACAGACTGCTTTAGACCTTACACACCTTCTTGATACCCAGCGAA	1800
Db	1741	TAAACCCAGAGCTACAGACTGCTTTAGACCTTACACACCTTCTTGATACCCAGCGAA	1800
Qy	1801	TTCAAACTGTTCACCCAGTCTCTTTCTGAGGGAACCGCTTCAAAATGTGACTGTGATGC	1860
Db	1801	TTCAAACTGTTCACCCAGTCTCTTTCTGAGGGAACCGCTTCAAAATGTGACTGTGATGC	1860
Qy	1861	CACCTGTTGCTGAAGTCTCAAACTATCAAGGACACACACCTTACCCAAAACATCTGTC	1920
Db	1861	CACCTGTTGCTGAAGTCTCAAACTATCAAGGACACACACCTTACCCAAAACATCTGTC	1920
Qy	1921	TGCACCAAAACCCATCTGTTCTCCATACGAGTCAATCAGTAAAGCTAGCAAGAGAGATC	1980
Db	1921	TGCACCAAAACCCATCTGTTCTCCATACGAGTCAATCAGTAAAGCTAGCAAGAGAGATC	1980
Qy	1981	AGCCAAGCTTGGCCAAAGGAATGAGGTGAAAAGAGTTATGAAAATGTTGATAGTGGGG	2040
Db	1981	AGCCAAGCTTGGCCAAAGGAATGAGGTGAAAAGAGTTATGAAAATGTTGATAGTGGGG	2040
Qy	2041	ATAAGAAAGAACAAGATTTACACTTCACCTATTACTGTTAGGAAAACAAAGAAAGATC	2100
Db	2041	ATAAGAAAGAACAAGATTTACACTTCACCTATTACTGTTAGGAAAACAAAGAAAGATC	2100

QY	2101	ANAGCGAAGGAATCTCGTATTCAAAGTTATTCTCTCCTCAAGCATTTAAATCTTTATGG	2160
DB	2101	AAAGCGAAGGAATCTCGTATTCAAAGTTATTCTCTCCTCAAGCATTTAAATCTTTATGG	2160
QY	2161	ACCAACATCTAGAAAATGTACTCAAAATCTCATCAGCAGCGTCTACATCGTAAAAACAAT	2220
DB	2161	ACCAACATCTAGAAAATGTACTCAAAATCTCATCAGCAGCGTCTACATCGTAAAAACAAT	2220
QY	2221	TAGAGAAATCAAAATGATCGCGGTGGATTATCTCAAGATGCCAGGATCAAAATCAGAAAGA	2280
DB	2221	TAGAGAAATCAAAATGATCGCGGTGGATTATCTCAAGATGCCAGGATCAAAATCAGAAAGA	2280
QY	2281	TGCTTTGCCAAAAGAATCTAAATACATCCGCTCTTAAAAGGCTAAAAATGGACAAGTCTA	2340
DB	2281	TGCTTTGCCAAAAGAATCTAAATACATCCGCTCTTAAAAGGCTAAAAATGGACAAGTCTA	2340
QY	2341	TGTTCTGTGAAGATAAGACACATAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGAA	2400
DB	2341	TGTTCTGTGAAGATAAGACACATAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGAA	2400
QY	2401	AAATAGATACTAAGCGTTTGTATGCAACAAAACCTCTCGAAAGAAAGATGTTCTTCTTC	2460
DB	2401	AAATAGATACTAAGCGTTTGTATGCAACAAAACCTCTCGAAAGAAAGATGTTCTTCTTC	2460
QY	2461	GAATCAAGTCGCTCATGTTAAAGGCTGAGAGAGATATCTGGCTGAAGCTGACAATGAAT	2520
DB	2461	GAATCAAGTCGCTCATGTTAAAGGCTGAGAGAGATATCTGGCTGAAGCTGACAATGAAT	2520
QY	2521	GGTAGTTGCTCTATATTATTTCATCCAGATAAGGACAAATTTATACTTTGTAATGGACT	2580
DB	2521	GGTAGTTGCTCTATATTATTTCATCCAGATAAGGACAAATTTATACTTTGTAATGGACT	2580
QY	2581	ACATCTCTGGGGTGATATGATGAGCGCTATTAAATAGAATGGGCATCTTCCAGAAAGTC	2640
DB	2581	ACATCTCTGGGGTGATATGATGAGCGCTATTAAATAGAATGGGCATCTTCCAGAAAGTC	2640
QY	2641	TGGCAGATTCTACATAGCAGAACTTACGTGTCAGTTGAAAGTGTTCAATAAAATGGGTT	2700
DB	2641	TGGCAGATTCTACATAGCAGAACTTACGTGTCAGTTGAAAGTGTTCAATAAAATGGGTT	2700
QY	2701	TTATTATAGAGATATTAAACCTGATAATATTGTTGATTGATCGTGATGGTCATATTAAAT	2760
DB	2701	TTATTATAGAGATATTAAACCTGATAATATTGTTGATTGATCGTGATGGTCATATTAAAT	2760
QY	2761	TGACTGACTTTGGCCTCTGCACCTGGCTTCAGATGGACACACAGATTTCAAGTACTATCAGA	2820
DB	2761	TGACTGACTTTGGCCTCTGCACCTGGCTTCAGATGGACACACAGATTTCAAGTACTATCAGA	2820
QY	2821	GTGGTGACCATCCAGCGCAAGATAGCATGGATTTCAGTAAATGGGGGATCCCTCAA	2880
DB	2821	GTGGTGACCATCCAGCGCAAGATAGCATGGATTTCAGTAAATGGGGGATCCCTCAA	2880
QY	2881	GCTGTGATGTGGAGACAGACTGAAGCCATTAGCGGGAGAGCTGCACGCCAGCACCGC	2940
DB	2881	GCTGTGATGTGGAGACAGACTGAAGCCATTAGCGGGAGAGCTGCACGCCAGCACCGC	2940
QY	2941	GATGCTAGCACATCTTTGGTTGGGACGCCAATATTATGCACCTGAAGTGTTCCTAC	3000
DB	2941	GATGCTAGCACATCTTTGGTTGGGACGCCAATATTATGCACCTGAAGTGTTCCTAC	3000
QY	3001	GAACAGGATACACAGTTGTGTGATTGGTGGAGTGTGGTGTATTCTTTTGAATGT	3060
DB	3001	GAACAGGATACACAGTTGTGTGATTGGTGGAGTGTGGTGTATTCTTTTGAATGT	3060
QY	3061	TGGTGGGACAACCTCTTCTTGGCCAAACACCATTAGAACACAAATGAAGGTTATCA	3120
DB	3061	TGGTGGGACAACCTCTTCTTGGCCAAACACCATTAGAACACAAATGAAGGTTATCA	3120
QY	3121	ACTGGCAACATCTCTTACATTCACCACACAGCTAAACTAGTCTCTGAAGCTTCTGATC	3180
DB	3121	ACTGGCAACATCTCTTACATTCACCACACAGCTAAACTAGTCTCTGAAGCTTCTGATC	3180
QY	3181	TTATTATTAAACTTTTGCCGAGGACCGGAAGATCGCTTAGGCAAGAAATGGTGTGATGAA	3240

527 GTTACCTCTCCACCCACCTCCATGAGAGGCCCCCTCCATGAGTACCACTCCACCT 118

QY 2103 GACCGAAGGAATCTCGTATTCAAAGTATTCCTCAAGCAATTTAAATCTTTATGGAG 2162
 DB GACCGAAGGAATCTCGTATTCAAAGTATTCCTCAAGCAATTTAAATCTTTATGGAG 2326
 QY 2163 CAACATGTAGAAATCTACTCAATCTCATCAGCAGCTCTACATCGTAAAAACAATTA 2222
 DB CAACATGTAGAAATCTACTCAATCTCATCAGCAGCTCTACATCGTAAAAACAATTA 2386
 QY 2223 GAGAATGAATGATGGGCTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATG 2282
 DB GAGAATGAATGATGGGCTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATG 2446
 QY 2283 CTTTGGCAAAAAGATCTAATTAATACATCCGCTTTAAAGGGCTTAAATGGACAACTATG 2342
 DB CTTTGGCAAAAAGATCTAATTAATACATCCGCTTTAAAGGGCTTAAATGGACAACTATG 2506
 QY 2343 TTTTGTGAAGATTAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAA 2402
 DB TTTTGTGAAGATTAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAA 2566
 QY 2403 GTAGATACTAAGGCTTTGTATGCAACAAAACTCTTCCGAAAGAAAGATGTTCTTCTCGA 2462
 DB GTAGATACTAAGGCTTTGTATGCAACAAAACTCTTCCGAAAGAAAGATGTTCTTCTCGA 2626
 QY 2463 AATCAAGTCGCTATGTTAAGGCTGAGAGAGATATCCTGGCTGGAAGCTGACAAATGAATGG 2522
 DB AATCAAGTCGCTATGTTAAGGCTGAGAGAGATATCCTGGCTGGAAGCTGACAAATGAATGG 2686
 QY 2523 GTAGTTTGGCTATATATTCATTCACAGATAAGGACAAATTTATATCTTTGTATGACATAC 2582
 DB GTAGTTTGGCTATATATTCATTCACAGATAAGGACAAATTTATATCTTTGTATGACATAC 2746
 QY 2583 ATTCTGGGGGTGATGATGATGAGCCTATTAAATAGAAATGGGCTCTTCCAGAAAGTCTG 2642
 DB ATTCTGGGGGTGATGATGATGAGCCTATTAAATAGAAATGGGCTCTTCCAGAAAGTCTG 2806
 QY 2643 GCACATTTCTACATAGCAGAACTTACCTGTGCGAGTTGAAAGTGTTCATAAATGGGTTTT 2702
 DB GCACATTTCTACATAGCAGAACTTACCTGTGCGAGTTGAAAGTGTTCATAAATGGGTTTT 2866
 QY 2703 ATTCATAGAGATATTAACCTTGATATATTTTGTATGATGCTGATGGTCAATTAATG 2762
 DB ATTCATAGAGATATTAACCTTGATATATTTTGTATGATGCTGATGGTCAATTAATG 2926
 QY 2763 ACTGACTTTGGCCTCTGCACTGGCTTCAGATGGACACAGATTTCTAAGTACTATCAGAGT 2822
 DB ACTGACTTTGGCCTCTGCACTGGCTTCAGATGGACACAGATTTCTAAGTACTATCAGAGT 2986
 QY 2823 GGTGACCATCCAGCAGATAGCATGATTTAGTAATGATGGGGGATCCCTCAAGC 2882
 DB GGTGACCATCCAGCAGATAGCATGATTTAGTAATGATGGGGGATCCCTCAAGC 3046
 QY 2883 TGTCTAGTGGAGACAGACTGAAGCATTAGACGGGAGAGCTGCACGCCAGCAGCAGCA 2942
 DB TGTCTAGTGGAGACAGACTGAAGCATTAGACGGGAGAGCTGCACGCCAGCAGCAGCA 3106
 QY 2943 TGTCTAGCACAATCTTTGGTTGGGACTCCCAATTTATATGCACTGAAGTGTGCTACGA 3002
 DB TGTCTAGCACAATCTTTGGTTGGGACTCCCAATTTATATGCACTGAAGTGTGCTACGA 3166
 QY 3003 ACAGGATACACAGTGTGTGATGTTGGTGGAGTGTGGTGTATCTTTTGAAGATG 3062
 DB ACAGGATACACAGTGTGTGATGTTGGTGGAGTGTGGTGTATCTTTTGAAGATG 3226
 QY 3063 GTGGACAACCTCTCTTTCTGGCAACAACACCATTAGAACACAAATGAAGTTATCAAC 3122
 DB GTGGACAACCTCTCTTTCTGGCAACAACACCATTAGAACACAAATGAAGTTATCAAC 3286
 QY 3123 TGGCAACATCTTTCATATTCACCAACAGCTAAACTAGTCTCTGAAGCTTCTGATCTT 3182
 DB TGGCAACATCTTTCATATTCACCAACAGCTAAACTAGTCTCTGAAGCTTCTGATCTT 3346
 QY 3183 ATTATTAACCTTTGCCGAGGAGCCGAGATCGCTTAGGCAAGATGTTGCTGATGAATA 3242

DB 3347 ATTATTAACCTTTGCCGAGGAGCCCAAGATCGCTTAGCAGAAATGGTCTGATGAATA 3406
 QY 3243 AAAGCTCATCCATTTTAAACAATTTGACCTTCCAGCTGAGCAGCAGTCTGCT 3302
 DB AAAGCTCATCCATTTTAAACAATTTGACCTTCCAGCTGAGCAGCAGTCTGCT 3466
 QY 3303 TCATACATTTCTAAATCAGACACCAACAGATACATCAAAATTTGATCCTGTTGATCCT 3362
 DB TCATACATTTCTAAATCAGACACCAACAGATACATCAAAATTTGATCCTGTTGATCCT 3526
 QY 3363 GATAAATTTATGGAGTATGATAACAGGAGAAATGTAATGACACTCTCAATGGATGG 3422
 DB GATAAATTTATGGAGTATGATAACAGGAGAAATGTAATGACACTCTCAATGGATGG 3586
 QY 3423 TATAAATTTGAAAGCATCTTCCGAGCTTATGTAATTTACCTTCCGAGGTTTTT 3482
 DB TATAAATTTGAAAGCATCTTCCGAGCTTATGTAATTTACCTTCCGAGGTTTTT 3646
 QY 3483 GATGACATGCTACCCATATAATTTATCCGAGCTTATGTAATTTGATGATACATTAATCA 3542
 DB GATGACATGCTACCCATATAATTTATCCGAGCTTATGTAATTTGATGATACATTAATCA 3706
 QY 3543 CAAGGCTCAGACGAGCAGTGGATGAGATGATCAAAACACAGGCTCAGAGATTAAAAAT 3602
 DB CAAGGCTCAGACGAGCAGTGGATGAGATGATCAAAACACAGGCTCAGAGATTAAAAAT 3766
 QY 3603 CGCATCTAGTATATGTTTAAACACACTAGTAAATTAATGTAATGAGGATTTGTAAGGG 3662
 DB CGCATCTAGTATATGTTTAAACACACTAGTAAATTAATGTAATGAGGATTTGTAAGGG 3826
 QY 3663 CCTGAATGCGAGGCTTTTGGAGTCTTGCAGATAAATTTATGCAAAATATGCAAGAGCTA 3722
 DB CCTGAATGCGAGGCTTTTGGAGTCTTGCAGATAAATTTATGCAAAATATGCAAGAGCTA 3886
 QY 3723 TATATGTGCTCTGTGTACAAATTTTATTTCTTAAATTTATGGGAAATCCTTTTAAAA 3782
 DB TATATGTGCTCTGTGTACAAATTTTATTTCTTAAATTTATGGGAAATCCTTTTAAAA 3946
 QY 3783 TGTAAATTTATCCAGCGTTTAAATCAGTATTTAGAAAAAATTTGTAAGGAAAGTA 3842
 DB TGTAAATTTATCCAGCGTTTAAATCAGTATTTAGAAAAAATTTGTAAGGAAAGTA 4006
 QY 3843 AATTATGAACCTGAATATATAGTCACTTCTTGGTACTTAAAGTACTTAAATAGTAGTG 3902
 DB AATTATGAACCTGAATATATAGTCACTTCTTGGTACTTAAAGTACTTAAATAGTAGTG 4066
 QY 3903 CTTTGTGTTAAAGGAGAAACCTGGTATCTATTTGTATATATGCTTAAATTTTAAAAA 3962
 DB CTTTGTGTTAAAGGAGAAACCTGGTATCTATTTGTATATATGCTTAAATTTTAAAAA 4126
 QY 3963 CAAGAGTTTTTGAATTTTTT 3984
 DB CAAGAGTTTTTGAATTTTTT 4148

RESULT 4
 AAZ51506
 ID AAZ51506 standard; DNA; 3213 BP.
 XX AAZ51506;
 XX
 DT 21-JUN-2000 (first entry)
 XX Mouse Lats (large tumour suppressor) DNA.
 XX
 KW Mouse; Lats; large tumour suppressor; cytostatic; vulnery;
 KW cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
 KW treatment; prevention; screening; cancer; skin; ovarian tumour;
 KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
 KW LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
 KW dysplasia; degenerative disorder; growth deficiency; physical trauma;
 KW hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX OS Mus musculus.
 XX PH Key Location/Qualifiers
 XX FT CDS 1..2889
 XX FT /*tag= a
 XX FT /product= "Lats protein"
 XX FT /partial
 XX PN WO200010602-A1.
 XX PD 02-MAR-2000.
 XX PF 18-AUG-1999; 98WO-US19068.
 XX PR 18-AUG-1998; 98US-0096996.
 XX PR 18-AUG-1998; 98US-0096997.
 XX PA (UYIA) UNIV YALE.
 XX XU T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;
 XX PI Turenchalk GS, Stewart RA;
 XX WPI: 2000-246496/21.
 XX P-PSDB: AAY70391.
 XX Use of lats proteins, complexes of lats and cdc2 for treating cancer
 XX that is refractory to treatment by standard chemotherapy and radiation
 XX therapy, and disorders associated with aberrant levels of cdc2 activity.
 XX Claim 44; Fig 13; 134pp: English.
 XX The present sequence is a DNA encoding mouse Lats (large tumour
 XX suppressor) protein which is a cell overproliferation inhibitor and a
 XX negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
 XX The present sequence is useful for treating cancer that is refractory
 XX to standard chemotherapy or radiation therapy such as hyperplasia,
 XX metaplasia, or dysplasia, and disorders associated with aberrant
 XX levels of cdc2 activity. Conditions treated by promoting cdc2 function
 XX include degenerative disorders, growth deficiencies, hypoproliferative
 XX disorders, physical trauma, lesions, and wounds: An animal model
 XX preferably a mouse, in which a lats gene has been disrupted by homologous
 XX recombination; e.g. a lats knock-out mouse, is used for screening
 XX compounds that can be used to treat or prevent cancer, particularly
 XX skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 XX associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 XX hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
 XX Sequence 3213 BP; 946 A; 764 C; 696 G; 807 T; 0 other;
 XX Query Match 62.38; Score 2483; DB 21; Length 3213;
 XX Best Local Similarity 87.48; Pred. No. 0;
 XX Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;
 XX 732 GTGAGCAATCAGTTAAACCGCAACAGAGCTGGAAAGTTCTTAAAGATCTCTAGTCTCT 791
 XX 1 GTGCAACATTCATTAACCGGAACAAAGCTGGAAGTTCTTAAAGAGTCTCTAGTCTCT 60
 XX 792 CAGAGGATGCGCGCCCACTAGGAGAAAGTGTGGCCTATATCTTGAGAGTCCCAACTCA 851
 XX 61 CAGAGACACGGCCCATCTCTAGGAGAAAAATGTGTTTATCTGTTGAAAGCCCAACTCA 120
 XX 852 CAGACAGATGAGNAGACCTTTGCTGGATCTGGTATATCAGCATTTGTTCAAGCTCAC 911
 XX 121 CAGCGGATGAGNAGACCTTTGCTGGATCCGGATTCAGCATTTGTTCAAGCTCAC 180
 XX 912 CCTAGCAACGACAGAGAGTGAACCCCAACACCTCAAGTAAAGAGTGTACTCTCT 971
 XX 181 CCAGCATGACAGAGAGTGAACCCCAACACCTCAAGTAAAGAGTGTACTCTCT 240
 XX 972 CCACCACCTCCAGAGGCGGACCTCCCGCTCCCAAGAGGTACAACCTCCACCTCCCGCTTCA 1031
 XX 1318 AACAGATTACAACCTTCACTTATCAGTCTGTTCCGAAAAACAAGAAAGATGAAGACGAAGA 1377

Db 241 CCACCACCTCCGAGAGGCGCAGACCCACCTCCCGAGGACACACTCCCGCTCCCGCTCC 300
 QY 1032 TGGGAACCAAACTCTCAAAAGCGCTATTTCTGAAACATGGAATAGTATCTCCCGA 1091
 Db 301 TGGGAACCAAACTCTCAAAAGCGCTACTCTGGGAACATGAGTAGTATCTCCCGA 360
 QY 1092 ATCTCTCTCTCCCACTCTGGGCGCATGGCAAGAGGCTATCTCTCCACCACTCTCAACACT 1151
 Db 361 ATCTCTCTCTCTCCCACTCTGGGCGGTGGCAGGAGGTACCTCTCCACCACTCTTACCCT 420
 QY 1152 TCCCCCATGAATCTCTCTTAATCAAGGACAGAGAGCATATGTTCTGTTCTGTTGGCAGA 1211
 Db 421 TCTCCCATGAATCTCTCTTAATCAAGGAGGTACAGGCTCAGAGGGCCATTAGTTCTGTTCCAGT 480
 QY 1212 CAACCAATCATCATGACAGTCTTAGCAAAATTTACTTTCCATCAGGAGGACCTTGAATG 1271
 Db 481 CAACCAATCATCATGACAGTCTTAGCAAAATTTACTTTTACACGAGGCGGCTTGAGGT 540
 QY 1272 CAGAAATGCTACTGGACAAACTGATTTTCATGATACACCAAAATTTGTTCTCTCTGGCACT 1331
 Db 541 CAGAAATGCTACTGGACAAACTGATTTTCATGATACACCAAAATTTGTTCTCTCTGGCACT 597
 QY 1332 GTGAATCGGAGCCACCACTCTCCATATCTCTGACAGAGTATGAGCAAAAGCCCTTCT 1391
 Db 598 GTGACTCGGAGCCACCACTCTCCATATCTCTGACCCAGGCTAATGAGCAAAAGCCCTCT 657
 QY 1392 GCTTTTACAAAGAGGCGATCTGCTCTCTCTGTCATATACAAATGGAAGTATCTCTCAG 1451
 Db 658 GCTTTTACAAAGAGGCGATCTGCTCTCTCTGTCATATCGCAATGGAAGTATCTCTCAG 717
 QY 1452 TCTATGATGTTGCCAAACAGAAATAGTATACATACATGGAACATATAACATAGTACTCT 1511
 Db 718 TCGATGATGTTGCCAAACAGGACAGTATACATGAGGCTTATATATTAATTAATGCTCT 777
 QY 1512 GGACTGCAAAACAAATGTCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1571
 Db 778 GGACTGCAAAACAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
 QY 1572 GGGCATGAAATCCCTTACATGACCACTTACATACCACTGAGTCAAGTCAAAATCTTTAATAC 1631
 Db 838 GGGCATGAAATCCCTTACATGACCACTTACATACCACTGAGTCAAGTCAAAATCTTTAATAC 897
 QY 1632 CCATTAGGAAATAGAGCAAGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1691
 Db 898 CCATTAGGAAATAGAGCAAGTCACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 957
 QY 1692 GCAATTAACAGGCTCTCTTATTAACAGGCTGTGAAAGTATGCGTGTATTAACACAGAG 1751
 Db 958 GCCATCACACCCGCTCTCTTATTAACAGGCTGTGAAAGTATGCGTGTATTAACACAGAG 1017
 QY 1752 CTACAGACTGCTTTAGCACTTACACACCTCTCTGATACCACTGATACCACTGATTTCAACTGTT 1811
 Db 1018 CTGACAGACTGCTTTAGCACTTACACACCTCTCTGATACCACTGATACCACTGATTTCAACTGTT 1077
 QY 1812 CAACCCAGTCTCTTTCTGAGGGAACCGCTTCAAAATGCTGATGCTGCTGCTGCTGCTGCTGCT 1871
 Db 1078 CAGCTTACCCCTTTCTGAGGCTACAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
 QY 1872 GAAGTCCAAACTATCAAGGAGGACCACTTACCACTTACCACTTACCACTTACCACTTACCACT 1931
 Db 1138 GAAGTCCAACTATCAAGGTCACCACTTATCACTTATCACTTATCACTTATCACTTATCACTTAT 1197
 QY 1932 CCATCTGTTCTCTCAGGATCAATCAGTAAAGCTTACCACTTACCACTTACCACTTACCACTTAC 1991
 Db 1198 CCATCTGTTCTCTCAGGATCAATCAGTAAAGCTTACCACTTACCACTTACCACTTACCACTTAC 1257
 QY 1992 CCCAAGGAGATGAGTGAAGGATTTATGAAATGTTAGTGGGATTAAGGAAAG 2051
 Db 1258 CCCAAGGAGATGAGTGAAGGATTTATGAAATGTTAGTGGGATTAAGGAAAG 1317
 QY 2052 AAACAGATTACAACCTTCACTTATTTAGTAAAGGAAAGGATGAAGGAGGAGG 2111
 Db 1318 AAACAGATTACAACCTTCACTTATTTAGTAAAGGAAAGGATGAAGGAGGAGG 1377

QY	3192	CTTTGCCGAGGACCGGAAGATCGCTTTAGCGAAGATGGTGTGATGAAATAAAGAGCTCAT	32351
Db	2458	CTGTGTCGAGGACCAAGACCGCCTCGCGAAGACGGTGTGATGAGATAAAGGCTCAT	2517
QY	3252	CCATTTTTTAAACAANTTGACATCTTCCAGTCAGCTGACAGACAGCAGTCTGCTTCATACATTT	3311
Db	2518	CCATTTTTTAAAGACCATCGATTCTCTAGTGTCTTGACACAGCAGTCTGCTTCATACATC	2577
QY	3312	CCTAAATATCACACACCCACAGACATACATAATTTTGATCTCTGTTGATCCTTGATAAATA	3371
Db	2578	CCTAAATATCACGATCCACACAGATACATCCAATTTTCGACCCCTGTTGATCCTTGATAAATTG	2637
QY	3372	TGGAGTATGATTAACGAGGAGAAATGTAAATGACACTCTCAATGGATGGTATATAAAT	3431
Db	2638	TGGAGCATGGCAGCGAGGAGGAAATATCAGTGACACTCTGAGCGGATGGTATATAAAT	2697
QY	3432	GGAAAGCATCTCGAACATGCTATGAATTTTACCTTCCGAAGGTTTTTTTGTGATGACAAT	3491
Db	2698	GGGAAGCACCCGAGCAGCGCTTCTATGAGTTTCACCTTTCGGAGGTTTTTTGATGACAAT	2757
QY	3492	GGCTACCCATATAATTATCCGAAGCCTATTTGAATATGAATATACATTAATTCCACAGGCTCA	3551
Db	2758	GGCTACCCATATAATTATCCGAAGCCTATTTGAATATGAATATACATTAATTCCACAGGCTCA	2817
QY	3552	GAGCAGCAGTCGGATGAAGATGATCAAAACACAGCGCTCAGAGATTTAAAAATTCGGGATCTA	3611
Db	2818	GAACAACAGTCTGATGAAGATGATCAACACACAAGCTCCGATGGAAACAACCGAGATCTA	2877
QY	3612	GTATATGTTTTAACACACTAGTAAATAAATGTAAATGAGGATTTGTAAAAGGCGCTGAATG	3671
Db	2878	GTGTATGTTTTAAATAACTAGGAGTCAATTTCTAA ----GAAATTTGCAAGAGGCGCTGAAGTG	2933
QY	3672	CGAGGTGTTTGGAGTCTCAGAGTAAATTTATGCAAAATATCACAGAGCTATATATGTGT	3731
Db	2934	CAGGGGTTTTTGAAGTTTTGAG ---AAAATTTGCAAAATCTCAGAG ----TTTGCTGT	2985
QY	3732	GCTCTGTGTACAATATTTTATTTTCCTAAATTTATGGGAAATCCTTTTTAAATGTTAAATTT	3791
Db	2986	GCTCTGTGTACAATATTTTATTTTCCTAAGTTATGGGAAATTTGTTTTAAAAATGTTAAATTT	3045
QY	3792	ATTCGAGCGGTTTTAAATCAGTATTTAGAAAAAATTTTATAGGAAACTAAATTTATGAA	3851
Db	3046	ATTCACCGCTTTTTAAATTCAGTAAATTTAGAAAAAATTTTATAGGAAAGTAATTTATGAA	3105
QY	3852	CTGAATATTAATAGTCAGTTCTTGGTACTTAAAGTACTTAAATTAAGTAG	3900
Db	3106	CTGACTATTAATAGTCAATCTTTGGTACTTTAAAGTACTTAAAGAGGAG	3154

RESULT 5	
AAT42119	
ID	AAT42119 standard; cDNA; 3213 BP.
XX AC	
XX AAT42119;	
XX	
DT	31-JAN-1997 (first entry)
XX	M-lats gene encoding large tumour suppressor.
DE	
XX	
XX	
KW	Mouse; m-lats gene; large tumour suppressor; fetal brain;
KW	protein-serine/threonine-kinase; cell proliferation; antisense;
KW	dominant-negative; cancer; degenerative disorder; trauma;
KW	growth deficiency; therapy; antitumour; vulnerrary; diagnostic;
KW	transgenic plant; transgenic animal; growth; senescence; ds.
XX	
OS	Mus musculus.
XX	
FH	Location/Qualifiers
FT	key 1..2889
FT	CDS /**tag= a
FT	/product= m-lats protein
XX	
PN	W09630402-A1.

XX 03-OCT-1996.
 XX PD
 XX 26-MAR-1996; 96MO-US04101.
 XX PF
 XX 27-MAR-1995; 95US-0411111.
 XX PR
 XX (UYVA) UNIV YALE.
 XX PA
 XX Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX PI
 XX WPI; 1996-455275/45.
 DR P-PSDB; AAW05179.
 XX
 PT New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX
 PS Disclosure: Page 126-130; 215pp; English.
 XX
 CC This sequence encodes a mouse large tumour suppressor m-lats protein,
 CC and has been isolated from a newborn mouse brain phage lambda-ZAP
 CC cDNA library using a 2.2-kb DNA probe from the Drosophila lats gene
 CC (AAT42117). A homologous mouse sequence has also been isolated
 CC (m-lats2, AAT42120). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 SQ Sequence 3213 BP; 946 A; 762 C; 697 G; 808 T; 0 other;

Query Match 62.2%; Score 2479.8; DB 17; Length 3213;
 Best Local Similarity 87.3%; Pred. No. 0;
 Matches 2767; Conservative 0; Mismatches 387; Indels 15; Gaps 4;

QY 732 GTGCAGCATCAGTTAACCGGAAACAGAGCTGGAAGGTTCTTAAGAAATCCTTAGTTCTCT 791
 DB 1 GTGCAGCATCAGTTAACCGGAAACAGAGCTGGAAGGTTCTTAAGAAATCCTTAGTTCTCT 60
 QY 792 CAGAGGCAATGGCCGCGCCTAGGAGAAAGTGTGGCTATCATCTCGAGAGTCCCAACTCA 851
 DB 61 CAGAGACAGGCGCCATCTCTAGAGAAATGTGTTTATCTGTTCTGAAAGCCCCCACTCA 120
 QY 852 CAGAGAGATGTAGGAGACCTTTGTCTGGATCTGGTATATCAGCATTTGTTCAAGCTCAC 911
 DB 121 CAGGGGATGTAGGAGACCTCTGTCTGGATCCGGCATTCGAGCATTTGCTCAAGCTCAC 180
 QY 912 CCTAGCAAGGACAGAGTGAACCCCGCCACACACCTCAAGTAGGAGTGTACTCTCT 971
 DB 181 CCAAGCAATGGACAGAGTGAACCCCGCCACACACCTCAAGTAGGAGTGTACTCTCT 240
 QY 972 CCACCACTCCCAAGAGGCGAGACTCCCTCCCAAGAGGTACAACTCCACCTCCCTTCA 1031
 DB 241 CCACCACTCCGAGAGGCGAGACCCACCTCCCGAGGACCACTCCCTCCCTCCCTCA 300
 QY 1032 TGGGAACCAACTCTCAAAAGGCGCTATCTTGAAACATGGAATACGTAATCTCCCGA 1091
 DB 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCCGA 360
 QY 1092 ATCTCTCTCTCCACCTGGGCGATGGCAAGAGGCTATCTCCACCACTCTCAACACT 1151
 DB 361 ATCTCTCTCTCTCCACCTGGGCGGTGGCAGAGGCGGTACCCCTCCACCACTCTTACCACT 420
 QY 1152 TCCCCCATCAATCTCTCAATCAAGGACAGAGGCGATTTGTTCTGTTCTGTTGSCAGA 1211
 DB 421 TCTCCCATGAATCCCTCTAGCAGGCTCAGAGGCGCATTTAGTTCTGTTCCAGTTGTGTA 480
 QY 1212 CAACCAATCATCATGCAGAGTTCTAGCAAAATTTAACTTTTCCATCAGGGAGACCTTGAATG 1271

DB 481 CAACCCATCATATGCAGAGTACTAGCAAAATTTAACTTTACACCGGCGAGCTGGAGTT 540
 QY 1272 CAGAATGGTACTGGACAAACTGATTCATGATACACCAAAATGTTGTCCTCTGCTGGCACT 1331
 DB 541 CAGAATGGTGGTGGTCAAGTCTGATTTTATCGTGCACAAA---TGTCCCACTGGTCT 597
 QY 1332 GTCAATCGGCGAGCCACCACTCCATATCTCTGACAGCAGCTAATGGAACAAAGCCCTCT 1391
 DB 598 GTGACTCGGCGAGCCACCACTCCATATCTCTGACCCAGCTAATGGAACAAAGCCCTCT 657
 QY 1392 GCTTTACAAACAGGGGATCTGCTGCTCTCTGATATACAAATGGAAGTATTCCTCAG 1451
 DB 658 GCTTTACAAACAGGGGCTTCTGCTGCTCCACCATCATTCGCCAATGGAACGTTCTCAG 717
 QY 1452 TCTATGATGCTGCCAAACAGAAATAGTCATAACAGTGAATATATACAAATGATGACCT 1511
 DB 718 TCGATGATGCTGCCAAACAGAACAGTACATACATGGAGCTTTATATATATATGTCCT 777
 QY 1512 GGACTGCAAAACAAATTTGGCTCAGTCATCTCTGCTCCAGCCAGTCATCCCGGAGCACT 1571
 DB 778 GGACTGCAAAACAGCTGGCCCGAGTCTCTGCTCTGCGAGTCATCCCGAGCGT 837
 QY 1572 GGGCATGAAATCCCTACATGGCAACCTAACATACAGTGAAGTCAAAATTTCTTTAATAAC 1631
 DB 838 GGGCATGAAATTTCTACATGGCAACCTAACATACAGTGAAGTCAAAATTTCTTTAATAAC 897
 QY 1632 CCATTTAGGAAATAGAGCAAGTCACTGCTGCTAATTTCTCAGCCCTCTGCTACACAGTCACT 1691
 DB 898 CCATTTAGGAAATAGAGCAAGTCACTGCTGCTAATTTCTCAGCCCTCTGCTACAGTCACT 957
 QY 1692 GCAATTTACACCAAGTCTCTATTTCAACAGCTGTGAAAGTATGCTGTATTTAAACCCAGAG 1751
 DB 958 GGCATCACACCCGCTCTTATTTCAACAGCCGCTGAAAGCATGGCGCTCTGAAACCCAGAG 1017
 QY 1752 CTACAGATGCTTTTAGCACCTACACACCTCTTCTGGATACCAAGCAATTTCAAACTGTT 1811
 DB 1018 CTGACAGTCTTTTAGCCCCCAACCCATCTCTTGGATGCCACAGCCAGTCTGAGACTGTT 1077
 QY 1812 CAACCCAGTCTTTCTCTGAGGAAACCGCTTCAAAATGCTGACTGTGATGCCACTGTTGCT 1871
 DB 1078 CAGCTTACCCCTTTTCTGAGGTTACAGTCTCAAGTGTGCTGCTGCTGCTGCTGCTGCT 1137
 QY 1872 GAAGTCTCAAACTATCAAGGACCCACCCACCTTACCCAAAACATCTGCTGCAACCAAAAC 1931
 DB 1138 GAAGTCTCAAGCTATCAAGTCTCAAGGCTCCACCCGCTTATCCAAAACATCTGCTACACCAAAAC 1197
 QY 1932 CCATCTGTTCTCCATACAGTCAATTCATAGCTTAGCAAGCTAGCAAGATCAGCCAGCTTG 1991
 DB 1198 CCATCTGTTCTCCATATGAGTCAAGTCAAGGCTTGAAGCTTGAAGATGAACAGCTAGCTTA 1257
 QY 1992 CCCAAGGAAGATGAGAGTCAAAAGGTTATGAAATGTTAGTAGTGGGATAAAGAAAG 2051
 DB 1258 CCCAAGGAAGATGATAGTGAAGAGTGGCGAGAGTGGTGTGCTGCTGGGATTAAGAAAG 1317
 QY 2052 AAACAGATTACAACTTCACTTATTTAGTAAAGAAACAAAGAAAGATGAAGGCGAAG 2111
 DB 1318 AAACAGATTACAACTTCACTTATCTGTTGCGAAAAACAAAGAAAGATGAAGACGAAGA 1377
 QY 2112 GAATCTGTTTCAAGATTATTTCTCTCAAGCATTTAAATTTCTTTATGAGCAACATGTA 2171
 DB 1378 GAGTCTCGGATTTCAGAGTTTACTCCCAAGGCTTTTAAAGTCTTCTTCTGAGCAGCAGTA 1437
 QY 2172 GAAATGTACTCAAACTCATCAGCAGCTCTACATCGTAAACAAATTAAGATGAATGA 2231
 DB 1438 GAGAAAGTCTTGAAGTCTCATCAGAGCTCTGCTGCGAAGAAAGCAGCTAGAAATGA 1497
 QY 2232 ATGATGCGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAA 2291
 DB 1498 ATGATGCGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAG 1557
 QY 2292 AAAGATCTTAATACATCGCTCTTAAAGGGCTTAAAGTGAAGTCTATGTTTGTGAAG 2351

Db 1558 AAAGAGCTAATATATTCGCTCTTAAAGGGCTAAATGGACAGCTATATGTTGTAAG 1617
 Qy 2352 ATAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTCTAGCAAGAAAGTAGATACT 2411
 Db 1618 ATAAAGACATAGGAATAGGAGCGTTTGGTGAAGTCTCTAGCAAGAAAGTAGATACT 1677
 Qy 2412 AAGGCTTTGATGCAACAAAACCTCTCGAAGAAAGATGTTCTTCTCGAAGTCAAGTC 2471
 Db 1678 AAGGCTTTGATGCAACAAAACCTCTCGAAGAAAGATGTTCTCTCGAAGTCAAGTC 1737
 Qy 2472 GCTCATGTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACATGAATGGTAGTTCGT 2531
 Db 1738 GCTCATGTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACATGAATGGTAGTTCGT 1797
 Qy 2532 CTATATTTATTCATTCCTCAAGATTAAGGACAAATTTATCTTTGTAATGGACTACATCTCTGGG 2591
 Db 1798 CTGTAATCTCTTCCAGGACAGGACAACTTGTAATTTGTAATGGACTACATCTCTGGG 1857
 Qy 2592 GGTGATATGATGAGCCTATTAATTAAGATGGCATCTTTCCAGAAAGTCTGGCAGCATTC 2651
 Db 1858 GGGATATGATGAGCCTATTAATTAAGATGGCATCTTTCCAGAAAGTCTGGCAGCATTC 1917
 Qy 2652 TACATAGCAGAACTTACCTGTGCAAGTGAAGTGTTCATAAAATGGGTTTATTCATAGA 2711
 Db 1918 TACATAGCAGAACTTACCTGTGCAAGTGAAGTGTTCATAAAATGGGTTTATTCATAGA 1977
 Qy 2712 CATATTAACCTGATATTAATTTGATTCATGCTGATGGTGCATATTAATTAACCTGACTTT 2771
 Db 1978 CATATTAACCTGATATTAATTTGATTCATGCTGATGGTGCATATTAATTAACCTGACTTT 2037
 Qy 2772 GGCCTCTGCACTGGCTTCAGATGGACACAGATTTCAAGTACTATCAGAGTGGTAGCCAT 2831
 Db 2038 GGCCTCTGCACTGGCTTCAGATGGACACAGATTTCAAGTACTATCAGAGTGGTAGCCAT 2097
 Qy 2832 CCACGGCAAGATACATGATTTTCAATTAATGAGGAGTCCCTCAAGCTGCTGATGT 2891
 Db 2098 CCACGGCAAGATACATGATTTTCAATTAATGAGGAGTCCCTCAAGCTGCTGATGT 2157
 Qy 2892 GGAGACAGACTGAAGCCATTAAGAGCGGAGAGCTGCAGCCAGCAGCAGGATGCTAGCA 2951
 Db 2158 GGGACAGACTGAAGCCATTAAGAGCGGAGAGCTGCAGCCAGCAGCAGGATGCTAGCC 2217
 Qy 2952 CATCTTTGGTGGGACTCCCAATTAATTAATGAGGAGTGGCTGAGCAAGGATAC 3011
 Db 2218 CATCTCTGTTGGGACTCCCAATTAATTAATGAGGAGTGGCTGAGCAAGGATAC 2277
 Qy 3012 ACACAGTGTGCTGATGGTGGAGTGGTCTTATTTTGAATGTTGGTGGGACAA 3071
 Db 2278 ACACAGTGTGCTGATGGTGGAGTGGTCTTATTTTGAATGTTGGTGGGACAA 2337
 Qy 3072 CCTCCTTTCTTGGCACAACACCAATTAAGAAACAAATGAAGGTTATCACTGGCAACA 3131
 Db 2338 CCTCCTTTCTTGGCACAACACCAATTAAGAAACAAATGAAGGTTATCACTGGCAACT 2397
 Qy 3132 TCTCTTCAATTCACCAAGCTAACTCAGTCCCTGAAAGCTTCTGATCTTATTTAA 3191
 Db 2398 TCTCTTCAATTCACCAAGCTAACTCAGTCCCTGAAAGCTTCTGATCTTATTTAA 2457
 Qy 3192 CTTTGGCGAGACCGGAGATCGCTTAGGCAAGATGGTCTGATGAATAAAGCTCAT 3251
 Db 2458 CTGTGTCGAGGACGAGAGACCGCTCGGCAAGAACGGTGTGATGAGTAAGAGCTCAT 2517
 Qy 3252 CCATTTTTTAAACAAATGACTTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATACATT 3311
 Db 2518 CCATTTTTTAAACAAATGACTTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATACATC 2577
 Qy 3312 CCTAAATATCACACACCAACAGATACATCAATTTTATGCTGTTGATCTGATTAATA 3371
 Db 2578 CCTAAATATCACACACCAACAGATACATCAATTTTATGCTGTTGATCTGATTAATA 2637
 Qy 3372 TGGAGTATGATACGAGGAGAAAGTAAATGACACTCTCAATGATGATGATTAATAAT 3431
 Db 2638 TGGAGTATGATACGAGGAGAAAGTAAATGACACTCTCAATGATGATGATTAATAAT 2697

RESULT 6

AA32981
 ID AAX32981 standard; DNA; 2442 BP.
 AC AAX32981;
 XX
 DT 30-JUN-1999 (first entry)
 XX
 DE Human warts protein encoding DNA.
 XX
 KW Human; warts protein; Drosophila; protein kinase; epithelial cell;
 KW cell regulation; cell proliferation; cell differentiation; tumour; ss.
 XX
 OS Homo sapiens.
 XX
 PN WO9915558-A1.
 XX
 PD 01-APR-1999.
 XX
 PF 24-AUG-1998; 98WO-JP03739.
 XX
 PR 24-SEP-1997; 97JP-0258689.
 XX
 PA (SUME) SUMITOMO ELECTRIC IND CO.
 XX
 PI Kishimoto T, Nagamine Y, Nishiyama Y, Niwa S, Saya H;
 DR WPI; 1999-244389/20.
 XX
 DR P-PSDB; AAY06890.
 XX
 PT Gene of vertebrates homologous to a Drosophila gene
 XX
 PS Claim 7; Page 44-52; 68pp; Japanese.
 XX

CC The invention relates to a human warts protein, homologous to Drosophila
 CC warts gene expression product. The warts expression product is a protein
 CC kinase involved in the regulation of proliferation and differentiation of
 CC epithelial cells and suppression of tumour formation. The gene, its
 CC expression product and antibodies are useful in the study of the

Best Local Similarity 99.9%; Pred. No. 0;				Matches 1; Indels 1; Gaps 1;			
Matches 1372; Conservative 0;				Mismatches 0;			
QY	2130	TATTCCTCCTCAAGCAATTAATTAATTTATGAGCAACAATGTAGAAAATGTACTCAAAATCT	2189				
Db	1	TATTCCTCCTCAAGCAATTAATTAATTTATGAGCAACAATGTAGAAAATGTACTCAAAATCT	60				
QY	2190	CATCAGCAGCGTCTACATCGTAAAAACAATTAGAGATGAATGATCGGGTTGGATTA	2249				
Db	61	CATCAGCAGCGTCTACATCGTAAAAACAATTAGAGATGAATGATCGGGTTGGATTA	120				
QY	2250	TCTCAAGATGCCAGGATCAATAGAGAAAGATGCTTTGCCAAAAGAAATCTAATTAATC	2309				
Db	121	TCTCAAGATGCCAGGATCAATAGAGAAAGATGCTTTGCCAAAAGAAATCTAATTAATC	180				
QY	2310	CGTCTTAAAGGGCTTAAATGACAAAGTCTATGTTTGAAGATAAAGACACTAGGAATA	2369				
Db	181	CGTCTTAAAGGGCTTAAATGACAAAGTCTATGTTTGAAGATAAAGACACTAGGAATA	240				
QY	2370	GGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATACTAAGGCTTTGTATGCAACA	2429				
Db	241	GGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATACTAAGGCTTTGTATGCAACA	300				
QY	2430	AAACTCTTCCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGGCTGAG	2489				
Db	301	AAACTCTTCCGAAAGAAAGATGTTCTTCTCGAAATCAAGTCGCTCATGTTAAGGCTGAG	360				
QY	2490	AGAGATATCTCGCTCAAGCTGACAAATGAATGGTAGTTGCTCTATATATTATTCATTCOA	2549				
Db	361	AGAGATATCTCGCTCAAGCTGACAAATGAATGGTAGTTGCTCTATATATTATTCATTCOA	420				
QY	2550	GATAAGGACAAATTTATACCTTTGTAATGAGTACATTCCTGGGGGTGATATGATGAGCCTA	2609				
Db	421	GATAAGGACAAATTTATACCTTTGTAATGAGTACATTCCTGGGGGTGATATGATGAGCCTA	480				
QY	2610	TTAATTAAGATGGCATCTTTCCAGAAAGTCTGGCAGATTTCTACATAGCAGAACTTACC	2669				
Db	481	TTAATTAAGATGGCATCTTTCCAGAAAGTCTGGCAGATTTCTACATAGCAGAACTTACC	540				
QY	2670	TGTGACGTTGAAAGTGTCTATAAAATGGGTTTATTCATAGAGATATTAAACCTGATAAT	2729				
Db	541	TGTGACGTTGAAAGTGTCTATAAAATGGGTTTATTCATAGAGATATTAAACCTGATAAT	600				
QY	2730	ATTTTGAATGATCGTGTGATGCTATATTAATTAAGTACGTTTGGCCTCTGCACTGGCTTC	2789				
Db	601	ATTTTGAATGATCGTGTGATGCTATATTAATTAAGTACGTTTGGCCTCTGCACTGGCTTC	660				
QY	2790	AGATGGACACACGATTTCTAAGTACTATCAGAGTGGTGACCATCCACGGCAAGATGACATG	2849				
Db	661	AGATGGACACACGATTTCTAAGTACTATCAGAGTGGTGACCATCCACGGCAAGATGACATG	720				
QY	2850	GATTTCACTAATGAATGGGGGATCCCTCAAGCTGCTCGATGTGGAGACAGACTGAAGCCA	2909				
Db	721	GATTTCACTAATGAATGGGGGATCCCTCAAGCTGCTCGATGTGGAGACAGACTGAAGCCA	780				
QY	2910	TTAGAGCGGAGAGCTGACCGCCAGCACCGATGCTTAGCACATTTCTTTGGTGGGACT	2969				
Db	781	TTAGAGCGGAGAGCTGACCGCCAGCACCGATGCTTAGCACATTTCTTTGGTGGGACT	840				
QY	2970	CCCAATTATATTGACCTGAAAGTGTGCTACGAACAGGATACACAGATTTCTGTGATTGG	3029				
Db	841	CCCAATTATATTGACCTGAAAGTGTGCTACGAACAGGATACACAGATTTCTGTGATTGG	900				
QY	3030	TGAGCTGTTGGTGTATTTCTTTTGAATGTTGGTGGGACAACTCTCTTCTGGCACAA	3089				
Db	901	TGAGCTGTTGGTGTATTTCTTTTGAATGTTGGTGGGACAACTCTCTTCTGGCACAA	960				
QY	3090	ACACCATTTAGAAACACAAATGAAGTGTATCACTGGGCAACATCTCTTCACTTCCACCA	3149				
Db	961	ACACCATTTAGAAACACAAATGAAGTGTATCACTGGGCAACATCTCTTCACTTCCACCA	1020				
QY	3150	CAAGCTAAACTCAGTCTGAACTTCTGATCTTATTTAATTTTAACTTTGCGGAGGACCCGAA	3209				

Db	2041	ACACAAATGAAGTTATCAACTGGCAACATCTCTTCAATTCACCAACCAAGCTAAACTC	2100				
QY	3162	AGTCTGAAGCTTCTGATCTTATTAATTAATTTGCGGAGGACCCGAGATGCTTAGGC	3221				
Db	2101	AGTCTGAAGCTTCTGATCTTATTAATTAATTTGCGGAGGACCCGAGATGCTTAGGC	2160				
QY	3222	AAGATGGTCTGATGAATAAAGCTCATCCATTTTAAAAACAATTTGACATTTCTCCAGT	3281				
Db	2161	AAGATGGTCTGATGAATAAAGCTCATCCATTTTAAAAACAATTTGATTTCTCCAGT	2220				
QY	3282	GACCTGACACAGCAGTCTGCTTATACATTCCTTAAATACACACCCACAGATACATCA	3341				
Db	2221	GACCTGACACAGCAGTCTGCTTATACATTCCTTAAATACACACCCACAGATACATCA	2280				
QY	3342	AATTTGATCTCTGTTGATCCTGATAAATATGAGTATGATGATGAGGAGGAGAAATGTA	3401				
Db	2281	AATTTGATCTCTGTTGATCCTGATAAATATGAGTATGATGATGAGGAGGAGAAATGTA	2340				
QY	3402	AATGACACTCTCAAT-GGATGGTATAAATGGAAGCATCTCTGAACATGCAATTTCTATGA	3460				
Db	2341	AATGACACTCTCAATGGGATGATAAATGGAAGCATCTCTGAACATGCAATTTCTATGA	2400				
QY	3461	ATTTACCTTCGAGAGTGTGTTGATGACATGGCTACCCATA	3502				
Db	2401	ATTTACCTTCGAGAGTGTGTTGATGACATGGCTACCCATA	2442				
RESULT 7							
AX32982	ID AX32982.standard; DNA; 1374 BP.						
XX	AC	AX32982;					
XX	DT	30-JUN-1999 (first entry)					
XX	DE	Human warts gene fragment encoding a partial peptide fragment.					
XX	KW	Human; warts protein; Drosophila; protein kinase; epithelial cell;					
XX	OS	cell regulation; cell proliferation; cell differentiation; tumour; ss.					
XX	PN	Homo sapiens.					
XX	PD	WO915558-A1.					
XX	PF	01-APR-1999.					
XX	PR	24-AUG-1998; 98WO-JP03739.					
XX	PA	24-SEP-1997; 97JP-0258689.					
XX	PI	(SUME) SUMITOMO ELECTRIC IND CO.					
XX	PT	Kishimoto T, Nagamine Y, Nishiyama Y, Niwa S, Saya H;					
XX	PS	WPI; 1999-244389/20.					
XX	PP	P-PSDB; AAY06891.					
XX	PP	Gene of vertebrates homologous to a Drosophila gene					
XX	PS	Disclosure; Page 52-56; 68pp; Japanese.					
XX	CC	The invention relates to a human warts protein, homologous to Drosophila					
XX	CC	warts gene expression product. The warts expression product is a protein					
XX	CC	kinase involved in the regulation of proliferation and differentiation of					
XX	CC	epithelial cells and suppression of tumour formation. The gene, its					
XX	CC	expression product and antibodies are useful in the study of the					
XX	CC	mechanisms of tumour development. The present sequence represents a					
XX	CC	human warts gene fragment encoding a partial peptide fragment.					
XX	SQ	Sequence 1374 BP; 442 A; 252 C; 288 G; 392 T; 0 other;					
Query Match			34.18;	Score 1360.4;	DB 20;	Length 1374;	

Seq	Sequence	1357 BP;	436 A;	254 C;	285 G;	382 T;	0 other;
	Query Match		31.0%;	Score	1233.2;	DB	22; Length 1357;
	Best Local Similarity		97.3%;	Pred. No.	1.1e-298;		
	Matches 1321;	Conservative	0;	Mismatches	23;	Indels	14; Gaps
QY	2276	AAAGATGCTTTGGCAAAAGAACTTAATTACATCCGCTCTTAAAGGGCTAAATGGACAA	2335				
Db	1	AAAGATGCTTTGGC-AAAGAAATCTAATTACATCCGCTCTTAAAGGGCTAAATGGACAA	59				
QY	2336	GTCTATCTTTTGTGAAGATAAAGACACATPAGGAATAGGAGCATTTGGTGGAGGTCCTCTPAGC	2395				
Db	60	GTCTATCTTTTGTGAAGATAAAGACACATPAGGAATAGGAGCATTTGGTGGAGGTCCTCTPAGC	119				
QY	2396	AAGAAAGTAGATACTAAGGCTTTGTATGTCAACAAAAAATCTTCGAAAGAAAGATGTTCT	2455				
Db	120	AAGAAAGTAGATACTAAGGCTTTGTATGTCAACAAAAAATCTTCGAAAGAAAGATGTTCT	179				
QY	2456	TCCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTCGCTGAAGCTGACAA	2515				
Db	180	TCCTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTCGCTGAAGCTGACAA	239				
QY	2516	TGAATGGGTAGTTCGCTATATATTATTCATTCACAGATAAGGACAAATTTATACTTTGTAAT	2575				
Db	240	TGAATGGGTAGTTCGCTATATATTATTCATTCACAGATAAGGACAAATTTATACTTTGTAAT	299				
QY	2576	GGACTACATTCCTGGGGGTGATATGATGAGCCTATTAAATTAGAATGGCATCTTTCCAGA	2635				
Db	300	GGACTACATTCCTGGGGGTGATATGATGAGCCTATTAAATTAGAATGGCATCTTTCCAGA	359				
QY	2636	AACTCTGGCACGATTCTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAAT	2695				
Db	360	AACTCTGGCACGATTCTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAAT	419				
QY	2696	GGGTTTTATTCATAGAGATATTAACCTGATATATTTTGATTTGATCGTGAGTGCATAT	2755				
Db	420	GGGTTTTATTCATAGAGATATTAACCTGATATATTTTGATTTGATCGTGAGTGCATAT	479				

QY 2756 TAAATTGACTGACCTTGGCCCTCTCCACTGGCTTCAGATGGACACACGATTTCTAAGTACTA 2815
 DB 480 TAAATTGACTGACCTTGGCCCTCTCCACTGGCTTCAGATGGACACACGATTTCTAAGTACTA 539
 QY 2816 TCAGATGGTGACCATCCAGGGCAAGATAGATGATTTTCAGTAATGAATGGGGGATCC 2875
 DB 540 TCAGATGGTGACCATCCAGGGCAAGATAGATGATTTTCAGTAATGAATGGGGGATCC 599
 QY 2876 CTCAGCTGTGATGTGGAGACACATGAAGCCATTAGAGCGGAGAGCTGCACGCCAGCA 2935
 DB 600 CTCAGCTGTGATGTGGAGACACATGAAGCCATTAGAGCGGAGAGCTGCACGCCAGCA 659
 QY 2936 CCACCGATGTAGCACATTTTGGTGGGACTCCCAATATATATTCACCTGAAGTGT 2995
 DB 660 CCACCGATGTAGCACATTTTGGTGGGACTCCCAATATATATTCACCTGAAGTGT 719
 QY 2996 GCTACGAACAGGATACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTTTGA 3055
 DB 720 GCTACGAACAGGATACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTTTGA 779
 QY 3056 AATGTTGGTGGGACAACCTCTTTTGGCACAAACACCATTAGAAACACAAATGAAGT 3115
 DB 780 AATGTTGGTGGGACAACCTCTTTTGGCACAAACACCATTAGAAACACAAATGAAGT 839
 QY 3116 TATCAACTGCAACATCTTTCACATTCACACAAAGCTAACTCAGTCCCTGAAGCTTC 3175
 DB 840 TATCAACTGCAACATCTTTCACATTCACACAAAGCTAACTCAGTCCCTGAAGCTTC 899
 QY 3176 TGATCTATTATTAATTTTGGAGGACCCGGAAGATCTTGGACCAAGTGTGCTGA 3235
 DB 900 TGATCTATTATTAATTTTGGAGGACCCGGAAGATCTTGGACCAAGTGTGCTGA 959
 QY 3236 TGAATTAAGCTATCCATTTTAAAAA---CAATTGACTTCT-CCAGTGAACCTGA--- 3288
 DB 960 TGAATTAAGCTATCCATTTTAAAAA---CAATTGACTTCTCCAGTGAACCTGA--- 1019
 QY 3289 ---GACAGCAGTCTGCTTATCATATCTTCTTAAATACACAC---ACCCACAGATACATCAA 3343
 DB 1020 CAGCAGCTGGCTTTTCAACAATTCCTTAAATACACAC---ACCCACAGATACATCAA 1079
 QY 3344 TTTTGATCC-TGTTGATCCGATAAATATGAGTGAATGATTAACAGGAGGAAAGATGAA 3402
 DB 1080 TTTTGATCCGTTGATCCGATAAATATGAGTGAATGATTAACAGGAGGAAAGATGAA 1139
 QY 3403 ATGACACTCTCAATGGATGATAAATGGAAGCATCTGACATGATTCATGAAT 3462
 DB 1140 ATGACACTCTCAATGGATGATAAATGGAAGCATCTGACATGATTCATGAAT 1199
 QY 3463 TTACCTTCGGAAGTGTGATGACATGCTACCCATATAATTTATCCGAGCTATTG 3522
 DB 1200 TTACCTTCGGAAGTGTGATGACATGCTACCCATATAATTTATCCGAGCTATTG 1259
 QY 3523 ATATGATATACATTAATTCACAGCTCAGAGCAGCTGCGATGAAGATGATCAAAACA 3582
 DB 1260 ATATGATATACATTAATTCACAGCTCAGAGCAGCTGCGATGAAGATGATCAAAACA 1319
 QY 3583 CAGGCTCAGAGATTAATAATCGCATCTAGTATATGTT 3620
 DB 1320 CAGGCTCAGAGATTAATAATCGCATCTAGTATATGTT 1357

RESULT 9

AAX32984

ID AAX32984 standard; DNA; 795 BP.

XX

AC AAX32984;

XX

DT 30-JUN-1999 (first entry)

XX

DE Human warts gene fragment encoding a partial peptide fragment.

XX

XX Human; warts protein; Drosophila; protein kinase; epithelial cell;

KW cell regulation; cell proliferation; cell differentiation; tumour; ss.
 XX Homo sapiens.
 OS WO9915558-A1.
 PN 01-APR-1999.
 PD 24-AUG-1998; 98WO-JP03739.
 PF 24-SEP-1997; 97JP-0258689.
 PR (SUME) SUMITOMO ELECTRIC IND CO.
 XX Kishimoto T, Nagamine Y, Nishiyama Y, Niwa S, Saya H;
 PI WPI; 1999-244389/20.
 DR P-PSDB; AAY06893.
 XX Gene of vertebrates homologous to a Drosophila gene
 PT Disclosure; Page 58-60; 68pp; Japanese.
 PS The invention relates to a human warts protein, homologous to Drosophila
 XX warts gene expression product. The warts expression product is a protein
 CC kinase involved in the regulation of proliferation and differentiation of
 CC epithelial cells and suppression of tumour formation. The gene, its
 CC expression product and antibodies are useful in the study of the
 CC mechanisms of tumour development. The present sequence represents a
 CC human warts gene fragment encoding a partial peptide fragment.
 XX Sequence 795 BP; 224 A; 228 C; 160 G; 183 T; 0 other;
 SQ

Query Match

Best Local Similarity 19.8%; Score 788.6; DB 20; Length 795;

Matches 791; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 786 GTTCCTCAGAGGCGATGCCCGCCACTAGGAGAAAGTGTGGCTATCATCTTGAGAGTCCC 845
 DB 1 GTTCCTCAGAGGCGATGCCCGCCACTAGGAGAAAGTGTGGCTATCATCTTGAGAGTCCC 60
 QY 846 AACTCAGACAGATGAGGAGACCTTTCTGTGATCTGTATATCAGCATTTGTTCAA 905
 DB 61 AACTCAGACAGATGAGGAGACCTTTCTGTGATCTGTATATCAGCATTTGTTCAA 120
 QY 906 GCTCACCTTAGCAACGACAGAGAGTGAACCCGCCACCCACCTCAAGTAAGAGTGT 965
 DB 121 GCTCACCTTAGCAACGACAGAGAGTGAACCCGCCACCCACCTCAAGTAAGAGTGT 180
 QY 966 ACTCCTCCACCTCCCAAGAGGCCAGACTCCCGCCCTCAAGAGGTACAACTCCACCTCC 1025
 DB 181 ACTCCTCCACCTCCCAAGAGGCCAGACTCCCGCCCTCAAGAGGTACAACTCCACCTCC 240
 QY 1026 CTTTCATGGGAACCAACTCTCAACAAAGCGTATCTGGAACATGGAATACGTAATC 1085
 DB 241 CTTTCATGGGAACCAACTCTCAACAAAGCGTATCTGGAACATGGAATACGTAATC 300
 QY 1086 TCCGGAATCTCTCTGTCCACCTGGGGCATGGCAAGGGGTATCTCCACACCTCTC 1145
 DB 301 TCCGGAATCTCTCTGTCCACCTGGGGCATGGCAAGGGGTATCTCCACACCTCTC 360
 QY 1146 AACACTTCCCGCATGAATCTCTTAATCAAGACAGAGGCGATTAGTTCTTCTCTGTT 1205
 DB 361 AACACTTCCCGCATGAATCTCTTAATCAAGACAGAGGCGATTAGTTCTTCTCTGTT 420
 QY 1206 GGCAGACAACCAATCATCATGAGAGTCTTACAAATTTAACTTTCCATCAGGGAGACT 1265
 DB 421 GGCAGACAACCAATCATCATGAGAGTCTTACAAATTTAACTTTCCATCAGGGAGACT 480
 QY 1266 GGAATGCAAGTGGTACTGGACAAACTGATTTTCATGATACACCAAAATTTCTCCCTGCT 1325
 DB 481 GGAATGCAAGTGGTACTGGACAAACTGATTTTCATGATACACCAAAATTTCTCCCTGCT 540

CC	proliferative disorders, degenerative disorders, trauma, growth
CC	deficiency, etc., and fragments of the gene may be used as diagnostic
CC	probes. A latex-inhibitor sequence may be expressed in a transgenic
CC	plant or farm animal to confer increased growth and inhibit
CC	senescence.
XX	
Sequence	3155 BP; 751 A; 924 C; 894 G; 596 T; 0 other;
Query Match	19.1%; Score 760; DB 17; Length 3155;
Best Local Similarity	63.7%; Pred. No. 6.8e-180;
Matches 1197; Conservative	0; Mismatches 665; Indels 18; Gaps
Qy	1653 CACTCTGTAATTTCTCAGCCCTTCTGTCACACAGTCACTGCAATTTACACAGCTCTTATT 1712
Db	985 CAACCTGAGCCCTCACTGCGCCGCCCAACACGGTCACCGCTGACGCCGCGCACATC 1044
Qy	1713 CAACAGCCTGTGAAGATATGCGTGATTTAAACACAGAGCTACAGACTGCTTTTAGCACCT 1772
Db	1045 CTTCAACCTGTGAAGACGCTGCGTGCTGCGGCCGAGCCCGACAGACGCGTGGGGCC 1104
Qy	1773 ACACACCTCTTCTGGATACACAGCCAAATTCACACTGTTTCAACCCAGTCTCTTCTCTGAG 1832
Db	1105 TCGCACCCCGCTGGGTGGTGGCGCCACAGCACCTGCCACTTGAGAGCTGGAGACGAG 1164
Qy	1833 GGAACCGCTTCAATGTGACTGTGATGCCACCTCTTGTCTGAAGCTTCCAAACTATCAAGGA 1892
Db	1165 GAGGCGAGCGCAGGCCACACACCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGGTGC 1224
Qy	1993 CCACCACACCTTACCCAAACATCTGCTGCACAAACCCATCTGTTCTTCTCATACAG 1952
Db	1225 CCACCGCTCCGATCCAAAGCACTTGTCTCTCCCGAGTAAGTCTGACAGTACAGCGTG 1284
Qy	1953 TCAATCAGTAGGCTAGCAAAAGGATCAGCAAGCTTGCCTCCCAAGAGATCAGAGTGAA 2012
Db	1285 GACCTGGACAGCTGTGCACCATGTGTCAGCAGAGTCTGCGAGGGGCGACTGATCTAGAC 1344
Qy	2013 AAGAGTTA-----TGAAAATGTTGATGTGGGATTAAGAAAGAAAGAACAG 2057
Db	1345 GGGAGTGACAAGAGCCACAAAGGTGCGAAGGAGACAAAGCTGGCAGAGACAAAAAGCAG 1404
Qy	2058 ATTACACTTCACCTATTACTGTTAGGAAAAACAGAAGATGAAGACCGAAGGGAATCT 2117
Db	1405 ATTCAGACCTCCCGGTGCGTGTCCCAAGAAATAGCAGAGATGAAGAGAAGAGAGTGCT 1464
Qy	2118 CGTATCAAAGTTATCTCCTCAAGCATTTAAATCTTTATGSGACACATGTAGAAAT 2177
Db	1465 CGCATCAAGAGTTACTCCCTTTATGCCCTTCAATTTCTCATGGAGCAACACGTGGAGAT 1524
Qy	2178 GTACTCAAACTCTCATCAGCAGCGCTTACATCTGTAATAAACAAATAGAGAATGAATGATG 2237
Db	1525 GTCATCAAAACCTACCAGCAGAGAGTCAAGCGGAGCTACACCTGGAGCAGGAATGCC 1584
Qy	2238 CGGTTGGATTTCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAGAA 2297
Db	1585 AAAGCTGGGCTCTGTGAGGCGCAGCAGCAGCAGATGAGGAAGATCCTCTACCAAGAGCAG 1644
Qy	2298 TCTAATTTACATCCGCTTTAAAGGGCTTAAATGGACAGTCTATCTTTGTCAAGATAAG 2357
Db	1645 TCTAAGTCAACCGCTGAAGAGGGCCAGATGGACAGTCCATTTGTGAAATCAAG 1704
Qy	2358 ACATAGGAATAGGACATTTGGTGAAGTCTCTGTAGCAAGAAAAGTATGATCTAAGGCT 2417
Db	1705 ACTTAGGCATCGTGCTTTGGGGAAGTGTGCTCGCTTGTGAAGCTGAGCAGCTACGCT 1764
Qy	2418 TTGTATGCACAAAACTCTTCAAGAGAAGATGTTCTTCTCGAAATCAAGTGCCTCAT 2477
Db	1765 CTGTACGCCATGAAGACTCTCAGGAAGAAGAGATGCTTGAACCCGAATCAAGTGGCCCAT 1824
Qy	2478 GTTAAGGCTCAGAGAGATATCTCTGGCTGAAGCTGACATGAATGGTAGTTCGTCATAT 2537
Db	1825 GTCANGGCTGAGAGGACATCTTGGCTGAAGCAGACAAATGAGTGGGTGGTCAAACTCTAC 1884
Qy	2538 TATTCAFTCCAAGATAGGACAAATTTATATCTTTGTAATGGACTACATTTCTCTGGGGTGAT 2597

Db 1885 TACTCCCTTCCAGGACAGGACGCTGCTACTTTGATGGACTACATACACGCGGGAT 1944
 Qy 2598 ATGATGAGCCATTAAATAGAGTGGGATCTTTCCACAAAGTCTGGACGATTTCTACATA 2657
 Db 1945 ATGATGAGCCCTGCTGATCAGATGGAGTCTTCCCTGAGCAGCCTGGCCGCTTCTACAT 2004
 Qy 2658 GCAGAACTTACCTGTCAGTGGAGTGGTTCATAAATGGTGTATTTATCATACAGATATP 2717
 Db 2005 GCAGCTTGGCCCTGGCCATTTGAAGTGTCCCAAGATGGCTTTATTCACCGGGACATC 2064
 Qy 2718 AAACCTGATATATTTGATGATGCGATGATGATGATGATGATGATGATGATGATGATG 2777
 Db 2065 AAGCCTGACACACATCTGACCTGATGGTGCATATTAAGCTGACAGATTTTGGCCCTC 2124
 Qy 2778 TGCACCTGGCTTCAGATGACACACAGATCTTAAGTACTATCAGAGTGGTACCATCCACGG 2837
 Db 2125 TGCACCTGGATTCAGTGGATCTCAATTCCTCAAGTACTACCAAGAGGAAACACATGAGA 2184
 Qy 2838 CAAGATAGCATGATTTCAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGAC 2897
 Db 2185 CAGGACAGCATGGAGCCGCTGACCTCTGGGACGATGTTTCCAACGTGCTGTGGAGAC 2244
 Qy 2898 AGACTGAAGCCATTAGAGCGGAGAGTGCACGCCACACAGGATGTCTAGACATTTCT 2957
 Db 2245 AGTTAAAGACCTTGGAGCAGAGGGCGCAGAAAGCAGACAGAGTGCCTGGCACATTTCT 2304
 Qy 2958 TTGTTGGGACTCCCAATATATTTGACCTGAACTGAGTGTCTGACGACAGGATACACAG 3017
 Db 2305 CTGTCGGGACACCAATATACATCGCTCCGGAGGTCTCTCGGAAAGGGTACACGAG 2364
 Qy 3018 TTCTGTGATGTGGAGTGTGGTGTATTTCTTTTGAATTTGTTGGGACAACTCTCT 3077
 Db 2365 CTCTGTGATGTGGAGTGTGGTGTATTTCTTTTGAATTTGTTGGGACAACTCTCT 2424
 Qy 3078 TTCTTGGGACCAACACCATTTAGAAACACAAATGAAGTGTATCACTGCAACATCTCTT 3137
 Db 2425 TTCTTGGGACCAACACCATTTAGAAACACAAATGAAGTGTATCACTGCAACATCTCTT 2484
 Qy 3138 CACATTCACCAACAGCTCACTGCTGAGCTTCTGATCTTATTTATTAATTAACCTTGC 3197
 Db 2485 CATATCCTTACGAGGTGAGGCTGAGGCTGAGGCGGAGAGCTCTATCAGCAAGCTGTGC 2544
 Qy 3198 CGAGGACCGGAGATCGCTTAGCAAGATGCTGCTGATGATGATGATGATGATGATGATG 3257
 Db 2545 TCGCGGCTGATGCTGCGGCTGGGAGGATGGGAGATGATGATGATGATGATGATGATGATG 2604
 Qy 3258 TTTAAACAATTTGACTTCTCCAGTGACCTGAGACAGAGTCTGCTTATCATATTTCTTAA 3317
 Db 2605 TTCAACACCATCGACTTTTCCGCTGACATCCGAAAGAGGCTGACACCTACGTCCTCCAC 2664
 Qy 3318 ATCACACACCCACAGATACATCAATTTGATCCTGTTGATCTGTGATGATGATGATGATG 3377
 Db 2665 ATCAGCCACCCCATGACACCTTCCAAATTTGACCGGTGATGATGATGATGATGATGATG 2724
 Qy 3378 GATGATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 3437
 Db 2725 GAGGCGAGC---GGAGAGGCGCCAGGCGCTGGACAGCTGGGCTCCCGCAGCAGCAAG 2781
 Qy 3438 CATCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3497
 Db 2782 CATCAGAGCAGCGGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATG 2841
 Qy 3498 CCATATATTTATCCGAGCC 3517
 Db 2842 CCCTTCCGCTGCCGAGGCC 2861

RESULT 11

AAZ51507

ID AAZ51507 standard; DNA: 3155 BP.

XX

AC AAZ51507;

XX

DT

XX

DE

XX

XX

KW

KW

KW

KW

KW

KW

XX

OS

XX

XX

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

FT

XX

XX

PN

XX

XX

PD

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

XX

21-JUN-2000 (first entry)

Mouse Lats2 (large tumour suppressor) DNA.

Mouse; Lats2: large tumour suppressor; cytostatic; vulnary;
 cell-overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
 treatment; prevention; screening; cancer; skin; ovarian tumour;
 soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
 LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
 dysplasia; degenerative disorder; growth deficiency; physical trauma;
 hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

Mus musculus.

Key Location/Qualifiers

CDS 1..2943

FT /tag= a

FT /product= "Lats2 protein"

FT /tag= b

FT /note= "This region is erroneously repeated in the

FT mouse Lats2 DNA sequence shown in figure 14"

FT 1921..2960

FT /tag= c

FT /note= "This region is missing in the mouse Lats2

FT DNA sequence shown in figure 14"

WO200010602-A1.

02-MAR-2000.

18-AUG-1999; 99WO-US19068.

18-AUG-1998; 98US-0096996.

18-AUG-1998; 98US-0096997.

(UYIA) UNIV YALE.

Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

Turenchalk GS, Stewart RA;

WPI: 2000-246496/21.

P-PSDB: AAY70392.

Use of lats proteins, complexes of lats and cdc2 for treating cancer
 that is refractory to treatment by standard chemotherapy and radiation
 therapy, and disorders associated with aberrant levels of cdc2 activity

Claim 44; Pages 112-117; 134pp; English.

The present sequence is a DNA encoding mouse Lats2 (large tumour
 suppressor) protein which is a cell overproliferation inhibitor and a
 negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
 The present sequence is useful for treating cancer that is refractory
 to standard chemotherapy or radiation therapy such as hyperplasia,
 metaplasia, or dysplasia, and disorders associated with aberrant
 levels of cdc2 activity. Conditions treated by promoting cdc2 function
 include degenerative disorders, growth deficiencies, hypoproliferative
 disorders, physical trauma, lesions, and wounds. An animal model
 preferably a mouse, in which a lats gene has been disrupted by homologous
 recombination, e.g. a lats knock-out mouse, is used for screening
 compounds that can be used to treat or prevent cancer, particularly
 skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;

Query Match

Best Local Similarity 19.1%; Score 760; DB 21; Length 3155;

Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

QY 1653 CACTCTGCTAAATCTCAGCCTTCTGCTACACAGTCACTGCAATTACACCAGTCTCTATT 1712
 Db 985 CAACTGAGCCCTACTGCGCCGCCACACAGGTACCGCGGTGAGCGCCGACACATC 1044
 QY 1713 CAACAGCCTGTGAAGAATGCGGTATTAAACACAGAGTACAGACTGCTTTAGCACCT 1772
 Db 1045 CTTACCCCTGTGAAGAGCGTGGTGTGCTGCGGCCGAGCCGACAGACCGCTGGGGGCC 1104
 QY 1773 ACACACCCCTCTTGGATACACAGCAATTCAACTGTTCAACCCAGTCTTTTCTGTAG 1832
 Db 1105 TCGCACCCCGCTGGGTGGCTGCGGCCACAGCACCTGCGCACTGAGAGCGCTGGAGACGA 1164
 QY 1833 GGAACCGCTTCAATATGTGATGTCACACCTGTGCTGAAAGTCCAACTATCAAGA 1892
 Db 1165 GAGGCGAGCGAGCCACACCCCTGGATGTGGACTATGGCGGCTCCGAGCGCAGGTGC 1224
 QY 1893 CCACACACCCCTACCCCAACATCTGTCACCAAAACCCATCTGTTCTCCATACAGAG 1952
 Db 1225 CCACCGCTTCCCTGATCCCAAGCACTTGTCTGCCAGTAAGTCTGAGCAGTACAGCGTG 1284
 QY 1953 TCAATCAGTAAGCCTAGCAAGAGGATCAGCAAGCTTCCCAAGGAAGATCAGAGTGA 2012
 Db 1285 GACCTGGACAGCTGTGCACAGTGTGCAGCAGAGTCTGCGAGGGGGCACTGATCTAG 1344
 QY 2013 AAGAGTTA-----TGAATAATGTTGATGTTGGGATAAAGAAAGAAACAG 2057
 Db 1345 GGGAGTGACAGAGCCACAAAGCTGGAGGGAGACAAAGCTGGCAGACAAAAGCAG 1404
 QY 2058 ATTCAACCTTACCCTATTACTGTTAGGAAACAAAGAGATGAAGAGCGAAGGAATCT 2117
 Db 1405 ATTCAAGCTTCCCGTGTCTGCCGAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
 QY 2118 CGTATTCAGATTCTCTCTCAAGCATTAATTTCTTTATGGCAACATGTAGAAAT 2177
 Db 1465 CGCATCAGAGTTACTCCCTTTATGCCCTTCAAAATCTTCATGGAGCAACACGTGGAG 1524
 QY 2178 GTACTCAATCTCATCAGCAGCGCTTACATCGTAAAAACAATTAGAATGAATGATG 2237
 Db 1525 GTATCAAAACCTACCAGCAAGAGTCCAGCGAGGTACAGCTGGAGCAGAAATGGCC 1584
 QY 2238 CGGTTGGATTATCAAGATCCGAGGATCAAAATGAGAAAGATGCTTTGGCAAAAGAA 2297
 Db 1585 AAAGCTGGGCTCTGTGAGCGCGAGGAGGAGATGAGGAGATCTCTACCAAGAGAG 1644
 QY 2298 TCTAATACATCCGCTTAAAGGGCTAAATGGACAGTCTATCTTTGTGAAGATAAG 2357
 Db 1645 TCTAATCAACCGCTGAAGAGGGCCAAAGATGGACAGTCTCATGTTGTGAAATCAAG 1704
 QY 2358 ACACAGGAATAGGAGCAATTTGGTGAAGTCTGCTAGCAAGAAAGTAGATCTAAGGCT 2417
 Db 1705 ACTCTAGGCATCGTGCCTTTGGGGAAGTGTGCTCGCTTGTGAAGTGGACACTCAC 1764
 QY 2418 TTGTATGCAAAAACCTCTCGAAGAAAGATGTTCTTCTCGAAATCAAGTCCGCTAT 2477
 Db 1765 CTGTAGCGCATGAAGCATCTCAGGAAGAGGATGCTCTGAACCGGAATCAAGTGGCC 1824
 QY 2478 GTTAAGGCTGAGAGATATCTGCTGAGCTGACATGAATGGTGTGCTCTATAT 2537
 Db 1825 GTCAAGCTGAGAGGAGATCTGCTGAAGCAGACATGAGTGGGTGAGTCAAACTCTAC 1884
 QY 2538 TATTCTTCAAGATAAGGACAATTTATTTGTGTAAGTACATCTTCTGGGGGTGAT 2597
 Db 1885 TACTCTTCCAGACAGGACAGCCTGTACTTTGTGTGAGTACATACAGCGCGGGAT 1944
 QY 2598 ATGATGAGCCTATTAATAGATGGGCATCTTTCCAGAAAGTCTGGCAGATTTACATA 2657
 Db 1945 ATGATGAGCCTGCTGATCAGGATGGAGTCTTCCCTGAGCACCTGGCCGCTTCTACAT 2004
 QY 2658 GCAGAACTTACCTGTGCAGTTCAGATGTTTCAATAAATGGGTTTATTATCAGATATT 2717
 Db 2005 GCAGAGTTGACCCCTGGCCATTGAAAGTGTCCACAAAGATGGGCTTTATCCACCGGACATC 2064

QY 2718 AAACCTGATAATATTTTGAATGATGCTGATGCTCATATTAATTAATGACTGACTTTGGCCTC 2777
 Db 2065 AAGCCTGACAACTACTCATCGACCTGGATGGTCTATATTAAGCTGACAGATTTGGCCTC 2124
 QY 2778 TGCACCTGGCTTCAGATGGACACAGGATTAAGTACTATACAGAGTGGTACCATCCAGG 2837
 Db 2125 TGCACCTGGATTCAGGTGGACTCACAATTCCAAGTACTACAGAAAGGAAACCATGAGA 2184
 QY 2838 CAAGTAGCATGATTCAGTAATGAATGGGGGGATCCCTCAAGCTGTCGATGTGGAGAC 2897
 Db 2185 CAGGACAGCATGGAGCCCGGTGACCTCTGGGACGATGTTTCCAAGTGTGCTGTGGAGAC 2244
 QY 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGACGCCAGCAGCAGCAGTGTCTAGCATTTCT 2957
 Db 2245 AGTTTAAGACCTTGGAGCAGAGGGCGGAGAGCAGCAGCAGGTTGCTGCACTTTCT 2304
 QY 2958 TTGGTTGGGACTCCCAATATATTCACCTGAAAGTGTGCTACGAACAGGATACACAG 3017
 Db 2305 CTTGTGGGACACCAAAATTACATCGCTCGGAGGTGCTTCCGCAAAAGGTTACACGAG 2364
 QY 3018 TTGTGTGATTTGCTGGAGTGTGGTATTCTTTTGAATGTTGGTGGGCAACCTCTCT 3077
 Db 2365 CTCTGTGACTGTGGAGCGTGGTGTGATTTCTTTGAGATGCTGGTGGGCGCGCT 2424
 QY 3078 TTTCTTGGCACAACACCACTTAGAAAACACAAATGAAGGTATCAACTGGCAACATCTCT 3137
 Db 2425 TTTCTTGGCCCCCACCACACAGAGCAGCTGAGAGTGTCACTTGGGAGAGCAGCTG 2484
 QY 3138 CACATTTCCACCAAGCTAAACTAGTCTGAAAGTCTTCTGATCTTTATTTAACTTTGC 3197
 Db 2485 CATATCCCTTACGAGTGTAGGCTCAGCGCTGAGGCCGAGACCTCATCACGAAGCTGTC 2544
 QY 3198 CGAGCACCAGAGATCGCTTAGGCAAGATGCTGTGATGAATAAAGCTCATCATTT 3257
 Db 2545 TCGCGGCTGACTGCGCGCTGGGCGAGGATGGGCGAGTGAACCTCAAGSCACACCGGTC 2604
 QY 3258 TTTAAACAAATTTGACTTCTCCAGTGCACCTGAGACAGCAGTCTGCTTACATTTCTTAAA 3317
 Db 2605 TTCAACACCATCGACTTTCCGCTGACATCCGAAGCAGGCTGACCTTACCTGCCACC 2664
 QY 3318 ATCAGACCCCAACAGATACATCAAAATTTGATCTGTTGATCTGCTGCTGCTGCTGAGT 3377
 Db 2665 ATCAGCCACCCCATGGACACCTTCCAAATTTGACCCGGTGGATGAAGAAAGCCCTGGCAC 2724
 QY 3378 GATGATAAGGAGGAAGAAATGTAATGACACTCTCAATGGATGCTATAAATAAGAAAG 3437
 Db 2725 GAGGCCAGC---GGAGAGAGCCCAAGGCTGGGACACGCTGGCCCTCCCCCAGCAG 2781
 QY 3438 CATCTGACATGCTATTCTATGAATTTACCTTCCGAAAGGTTTTTTGATGACAAATGGCTAC 3497
 Db 2782 CATCCAGAGCAGCCTTCTATGAGTTTCACTTCCGAGGTTCTTCGATGACAAACGCTAT 2841
 QY 3498 CCATATATTTATCCGAAGCC 3517
 Db 2842 CCCTTCCGGTGCCCGAAGCC 2861

RESULT 12

AA87397
 ID AA87397 standard; cDNA; 5276 BP.

XX AA87397;

AC AA87397;

DT 08-OCT-1999 (first entry)

XX Human WART2 cDNA.

XX WART2; hWART2; WART orthologue; human; signal transduction;

KW protein kinase; cancer; tumour; diagnosis; therapy; ss.

XX Homo sapiens.

XX OS Homo sapiens.

XX FH Key Location/Qualifiers

FT 5'UTR 1..374
 FT CDS /*tag= a
 FT 375..3641
 FT /*tag= b
 FT 3642..5276
 FT /*tag= c
 XX
 PN W09937787-A2.
 XX
 XX 29-JUL-1999.
 XX
 XX 20-JAN-1999; 99WO-US01145.
 XX
 XX 21-JAN-1998; 98US-0072023.
 XX
 XX (SUGEN-) SUGEN INC.
 XX
 XX Flanagan P, Plowman GD;
 XX
 XX WPI; 1999-458698/38.
 XX
 XX P-PSDB; AAY06527.
 XX
 XX New nucleic acid encoding human orthologs of Drosophila WART
 XX proteins, used to identify specific modulators for treating cancer
 XX or for diagnosis
 XX
 XX Claim 1; Page 122-123; 137pp; English.
 XX
 XX This is the nucleotide sequence of a cDNA clone coding for a human
 XX orthologue, i.e. hWART2 (see AAY06527), of Drosophila non-receptor
 XX serine/threonine kinase WART. hWART2 cDNA was isolated from a
 XX human bone marrow cDNA library using a PCR fragment of hWART1 as
 XX probe. hWART2 is consistently expressed in human tumour cells
 XX lines, except for most of the colon cancer lines examined. This
 XX overexpression in tumour cells versus normal tissues may provide a
 XX target for oncology drug development. Nucleic acids encoding
 XX full-length hWART2 and hWART2 polypeptides lacking one or more of
 XX amino acid segments 1-33, 43-139, 342-466, 467-480, 514-518,
 XX and 974-1048, or lacking one or more of the N-terminal domain,
 XX catalytic domain, or C-terminal domain are claimed, as well as
 XX hWART2 sequences (AAX87396), hWART2 and hWART1 polypeptides,
 XX antibodies, a method for identifying modulators of hWART function,
 XX and use of such modulator compounds to treat an abnormal condition
 XX involving hWART signal transduction, especially cancer. Probes for
 XX detection of hWART nucleic acids are also claimed.
 XX
 XX Sequence 5276 BP; 1348 A; 1407 C; 1345 G; 1176 T; 0 other;
 XX
 XX Query Match 18.7%; Score 744.4; DB 20; Length 5276;
 XX Best Local Similarity 66.3%; Pred. No. 7e-176;
 XX Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps 3;
 QY 1872 GAAGCTCCAACTATCAAGGACCAACACCCCTACCAAAACATCTGCT-----GCACC 1926
 DB 1893 GGAGGCCGACAGCCGAGGAGTGCCTCCGCTTACCGGAGGACCTGCTGCTGCGGAGC 1952
 QY 1927 AAAACCCATCTGTTCTCCATACGAGTCAATCACTAGAGCTAGCAAGAGGATCAG----1982
 DB 1953 AAGTCGAGCAGTACGACCTGGACAGCCTGTGCGCAGCATGGAGCAGACCTCCGTGGC 2012
 QY 1983 ---CCAAAGCTTGCACCAAGGAGATGAGAGTCAAAAGAGTTATGAAATGTTGATAGTGGG 2039
 DB 2013 GGCCCCAACGAGCCGAGGGGCGGACAAAGAGCGGCAAGAGCGGCAAGAGGCAAGAGC 2072
 QY 2040 GATTAAGAAAGAAACAGATTACAACTTACCTATTACTGTAGGAAAAACAGAAAGAT 2099
 DB 2073 GGAAGGATAAAAAGCAGATTACAGACCTCTCCCGTTCCTCCGCAAAACACAGCAGAGAC 2132
 QY 2100 GAAGCAGCAAGGGAATCTCTGATTCAAGTATTCTCTCAAGCATTTAAATCTTTATG 2159
 DB 2133 GAAGAGAGAGAGTACGACATCAAGAGTACTCGCCATACGCTTTAAGTCTTCATG 2192
 QY 2160 GAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGCGTCTATCATGTAACAAACAA 2219

Db 2193 GAGCAGCAGCTGGAGAATGTCTATCAAAACCTTACCAGCAGAGAGTTAAACCGAGGCTGCAG 2252
 QY 2220 TTACAGATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAG 2279
 Db 2253 CTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGAGCAGATCGGAAG 2312
 QY 2280 ATGCTTTGCCAAAGAAATCTAATTTACATCGTCTTAAAGGGCTAAATGAGCAAGTCT 2339
 Db 2313 ATCTCTACCAAGAAAGTCTAATTTACAAAGGTTAAAGAGGGCCAAAGATGGCAAGTCT 2372
 QY 2340 ATGTTTGTGAAGATAAAGACACTAGGAATAGGACATTTGGTGAAGTCTGTCTAGCAAGA 2399
 Db 2373 ATGTTTGTCAAGATCAAAACCTCGGGATCGTCCCTTTGGAGAAGTGTGCTTGTCTGT 2432
 QY 2400 AAGTAGATACTAAGGCTTTGTATGCAACAAAACCTCTTCGAAAAGAAAGATGTTCTTCT 2459
 Db 2433 AAGTGGACACTCAGCCCTGTACGCCATGAAGACCTTAAGAAAAAGGATGTCTCTGAAC 2492
 QY 2460 CGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTTGGCTGAAGCTGACAAATGAA 2519
 Db 2493 CGGAATCAGTGGGCCACGCTCAAGCGCGAGAGGACATCTTGGCCGAGGACAGACATGAG 2552
 QY 2520 TGGTAGTTCGTCTATATTATTCATTCAGAGATAAGGACAAATTTATCTTTGTAATGGAC 2579
 Db 2553 TGGGTGTCAAACTCTACTCTCTCTCCAGACAAAGACAGCCTGTACTTTGTGATGGAC 2612
 QY 2580 TACATTCCTGGGGGTGATATGATGAGCCTATTATTAGATGGGCATCTTTCCAGAAAGT 2639
 Db 2613 TACATCTCTGGTGGGACATGATGAGCCTGCTGATCGGATGGAGGTCTTCCCTGAGCAC 2672
 QY 2640 CTGCGACGATTCTACATAGCAGAACTTACCTGTGTCAGTGTGAAGTGTTCATAAAATGGT 2699
 Db 2673 CTGCCCCGTTCTACATCGCAGACCTGACTTTGGCCATTTAGAGGTGCCAAGATGGC 2732
 QY 2700 TTTATTCATAGATATTTAAACCTGATATTTTGTGATCTGATGCTGATGCTATTTAAA 2759
 Db 2733 TTTATTCACCCGAGACATCAAGCCTGATAACTTTTGTATGATGCTGATGCTATTTAAA 2792
 QY 2760 TTGACTACCTTTGGCTCTGCACTGGCTTCAGATGGACACAGATTTCTAAGTACTATCAG 2819
 Db 2793 CTCACAGATTTGGCCCTCTGCACTGGCTTCAGTGGTCTCACAATTTCCAAATATTACAG 2852
 QY 2820 AGTGGTCACCATCCACGGCAAGATGATGATTTCAATGAATGGGGGATCCCTCA 2879
 Db 2853 AAGGGAGACCATGTCACAGAGGACAGCATGGAGCCAGCAGCCTCTGGGATGATGTCT 2912
 QY 2880 AGCTGTCCATGTGGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCACCCAGCAGCAG 2939
 Db 2913 AACTGTGCGTGTGGGACAGGCTGAAGACCTTAGAGCAGAGGGCGGGAAGCAGCAGCAG 2972
 QY 2940 CGATCTACACATCTTTGGTGGGACTCCCAATTTATTTGCACTGAAGTGTGCTA 2999
 Db 2973 AGTGGCTGACATCTTCTGTTGGGACTCCAACTTACATCGCAGCCGAGGTGCTCTC 3032
 QY 3000 CGAACAGGATACACAGTGTGTGATTGTGGAGTCTGGTGTATTCTTTTGAAGATG 3059
 Db 3033 CGCAAGGGTACACTCAACTCTGTACTGTTGGAGTGTGGAGTGTCTCTTCGAGATG 3092
 QY 3060 TTGGTGGGCAACCTCTCTTTTGGCAACAAACACCATTTAGAAAACACAAATGAAGTTATC 3119
 Db 3093 CTGGTGGGCGAGCGCCCTTTTGGCACCCTACTCCCAACAAACCCAGCTGAAGGTGATC 3152
 QY 3120 AACTGGCAACATCTCTTACATTTCCACACAGTCAAACTAGTCTCTGAACTTCTGAT 3179
 Db 3153 AACTGGGAGAACCGCTTCCACATTTCCAGCCAGGTGAAGCTGAGCCCTGAGCCGAGGAC 3212
 QY 3180 CTATTATTAAACTTTTCCCGAGGACCGGAAGATCGCTTAGCAAGATGTGCTGTATGAA 3239
 Db 3213 CTATCACCAAGCTGTGCTCTCGGAGACACCGCTGGGGCGGATATGGGGCGGATGAC 3272
 QY 3240 ATAAAAGCTCATCCATTTTAAAAACAAATTTGACTTCTTCCAGTGAACCTGAGACAGTCT 3299

Db 3273 CTGAAGCCGCCCCCTTCTTTACGGCCATTGACTTCTCCAGTGACATCCGGAAGCAGCCA 3332
 QY 3300 GCTTCATACATCTCTAAATACACACCCCAACAGATACATCAAAATTTTGATCTGTGTGAT 3359
 Db 3333 GCCCCCTACGTTCCACCATCAGCCACCCCATGACACCTCGGAATTCGACCCCGTAGAT 3392
 QY 3360 CCGTATAAATTTATGGAGTGATGATAAGCAGGAGAAATAATGTAATGACACTCTCAATGGA 3419
 Db 3393 GAAGAAGGCCCTTGGACGATGCCAGC---GAAGGTAGCACCAGGCCCTGGGACACACTC 3449
 QY 3420 TGGTATAAATGAAGATCTCGACATCGATTCATGCAATTTACCTTCCGAAGTTT 3479
 Db 3450 ACCTCGCCCAATACACGATCTCGACGACGATTTTACGAATTCACCTTCCGAAGTTT 3509
 QY 3480 TTTGATGACATGCTACCCATATATTTATCCGAAGCCTATTGAATATGAATACATTAAT 3539
 Db 3510 TTTGATGACATGCTACCCCTTCGATGCCAAGCCTTCAGGAGCAGGAAGCTTCACAG 3569
 QY 3540 TCACAAGGCTCAGA 3553
 Db 3570 GCTGAGAGCTCAGA 3583

RESULT 13

AA59129
 ID AA59129 standard; DNA; 5486 BP.

XX AA59129;

AC

XX 07-NOV-2000 (first entry)

DE DNA encoding a tumour suppressor protein hGHITS1.

XX Human; growth hormone inhibited tumour suppressor protein; hGHITS;
 KW antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 KW diabetic nephropathy; cardiopathy; tumour; breast cancer;
 KW renal adenocarcinoma; colorectal cancer; leukaemia; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
 FT CDS 387..3653
 FT /tag= a
 FT /product= "tumour suppressor protein"

XX EP1022333-Al:

XX 26-JUL-2000.

XX 07-OCT-1999; 99BP-0119199.

XX 25-JAN-1999; 99JP-0016223.

XX (JCRP-) JCR PHARM CO LTD.

XX Koga J, Kono K, Zolotar'ov FN;

XX WPI: 2000-516013/47.

XX P-PSDB; AAB07633.

XX New human growth hormone inhibited tumour suppressor genes 1 and 2 for

XX diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic

XX nephropathy or cardiopathy.

XX Claim 1; Page 18-28; 59pp; English.

XX The present sequence encodes a human growth hormone inhibited tumour

XX suppressor (hGHITS) protein. The hGHITS proteins can be used in an

XX antineoplastic pharmaceutical preparation. Probes for the hGHITS

XX DNA sequences can be used in diagnostic pharmaceutical preparations.

XX The diagnostic pharmaceutical preparations can be used for examining

XX expression of hGHITS genes in dwarfism, gigantism, acromegaly,

XX angiopathy, diabetic nephropathy or cardiopathy, or in malignant

CC tumours including breast cancer, renal adenocarcinoma, colorectal
 CC cancer, and leukaemia. Antibodies against the proteins can be used in
 CC a diagnostic pharmaceutical preparation for examining expression of
 CC a tumour suppressor gene.

XX
 SQ Sequence 5486 BP; 1378 A; 1452 C; 1390 G; 1266 T; 0 other;

Query Match 18.7%; Score 744.4; DB 21; Length 5486;
 Best Local Similarity 66.3%; Pred No. 7.2e-176;
 Matches 1123; Conservative 0; Mismatches 556; Indels 15; Gaps 3;

QY 1872 GAAGCTCCAACTATCAAGGACCCACCCCTACCCCAAAACATCTGCT-----GCAC 1926
 Db 1905 GGAGGCCAGACCGGAGGTGCCCGCTCCGCCCTACCCGAAGCACCTGCTGCTGCAGC 1964
 QY 1927 AAACCCATCTGTTCTCCATACGAGTCAATCAGTAGCCTAGCAAGAGGATCAG---- 1982
 Db 1965 AGTCGGAGCAGTACGAGCCTGAGCCTGTCGCGAGGCTGAGGAGGAGGAGCCTCCGTCG 2024
 QY 1983 ---CCAAAGTTGCCCAAGGAAGATGAGAGTGAAGAGATTATGAAATGTTGATAGTGG 2039
 Db 2025 GGCCCAACGAGCGCGGCGGCGGACAGAGCCGCAAGGCCCAAGGGGACAAAGGC 2084
 QY 2040 GATAAGAAAAAGAACAGATTACAACTTCACTTACTTCTTAGGAAAAACAAGAAAGAT 2099
 Db 2085 GGAAAGGATAAAAAAGCAGATTACAGCTCTCCCGTTCCCGTCCGCAAAACACAGCAGAC 2144
 QY 2100 GAAGCGGAAGGAATCTCGTATTCAAGATTATCTCTCAAGCATTTAAATCTTTATG 2159
 Db 2145 GAAGAGAGAGAGAGTACGATCAAGAGCTACTCGCCATACGCTTAAAGTCTTTCTATG 2204
 QY 2160 GAGCAACATGTAGAAAATGTACTCAAACTCTATCAGCAGCGTCTACATCGTAAAAACAA 2219
 Db 2205 GACACACGCTGGAGAAATGTATCAAAACCTACCCAGAGAGTTAACCGAGGTGCGAG 2264
 QY 2220 TTAGAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAG 2279
 Db 2265 CTGGCAAGAAATGCCAAAGCTGCTGTGAAGTGGAGGAGCAGATGCGGAAG 2324
 QY 2280 ATGCTTTGCCAAAGAAATCTAATTAATCTATCCGCTTTAAAGGCTAAAATGCAAGTCT 2339
 Db 2325 ATCTCTACCAAGAAAGAGTCTAATTAACAAGGTTAAAGAGGCGCAAGATGGAACAGTCT 2384
 QY 2340 ATGTTTGTGAAGATAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTGTCTAGCAAG 2399
 Db 2385 ATGTTTGTCAAGATCAAAACCTTGGGATCGGTGCGCTTTGGAGAGTGTGCTTGTGT 2444
 QY 2400 AAAGTAGATACTAAGCTTTGTATGCAACAAACAACTCTTCGAAGAAAGATGTTCTTCTT 2459
 Db 2445 AAGGTGGACACTCAGCCCTGTACGCCATGAAGACCTTAAGGAAAAAGATGTCTTGAAC 2504
 QY 2460 CGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACATGAA 2519
 Db 2505 CGGAATCAGGTGGCCCGCAGCGTCAAGCGGAGGAGATCCTGGCCGAGGACACAAATAG 2564
 QY 2520 TGGTAGTTCGCTCTATATTATTATTCATTCGAAGATAAGGACAAATTTATCTTTGTAATGAC 2579
 Db 2565 TGGGTGTCAAACTCTACTACTCTTCCAAAGACAAAGACAGCTGTACTTGTGATGGAC 2624
 QY 2580 TACATTCCTGGGGGTGATATGATGAGCCCTATTAAATTAAGATGGGCATCTTCCAGAAAT 2639
 Db 2625 TACATCCCTTGGGGACATGATGAGCCCTGTATCCGATGGAGTCTTCCCTGAGCAC 2684
 QY 2640 CTGGCAGCATTTACATAGCAGACACTTACCTGTGCAAGTGAAGAGTCTTCAATAAATGGT 2699
 Db 2685 CTGGCCCGGTTTCTACATCGCAGAGCTGACTTTGGCCATTCGAGAGTTCACCAAGATGGC 2744
 QY 2700 TTATTATAGAGATATTAAACCTGATAATATTTTATTGATGCTGATGCTGATATATAA 2759
 Db 2745 TTCATCCCGGAGACATCAAGCCTGATACATTTTGTATAGATCTGGATGCTGATATAA 2804
 QY 2760 TTGACTGACTTTGGCCCTCTGCACTGGCTTCAGATGGACACAGCATTTCTAAGTACTATCAG 2819

Db 2325 ATCCCTTACAGAAAGAGTCTAATTAACAAGGTTAAAGAGGCCAAGATGGACAAAGTCT 2384
QY 2340 ATGTTTGTGAAGATAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTGCTTAGCAAGA 2399
Db 2385 ATGTTTGTGAAGATCAAAACCCCTGGGGATCGGTGCGCTTTGGAGAAGTGGCCCTTGTGT 2444
QY 2400 AAGTATGATCTAAGCTTTGTATGCAACAAACAACTCTTCGAAGAAGAAATGTTCTTCT 2459
Db 2445 AAGTGGACACTCACGCCCTGTAGCCATGAAGACCCCTAAGGAAAAAGGATGCTCTGAAC 2504
QY 2460 CGAAATCAAGTCCCTCATGTTAAGGCTGAGAGAGATATCTGCTGAGCTGACATGAA 2519
Db 2505 CGAAATCAGTGGCCCGTCAAGCGGAGAGGGAGATCTGCGCCGAGCGACAAATGAG 2564
QY 2520 TGGGTAGTTCGTCTATATTATTCATTCACAAAGATAAGACAAATTTATCTTTGTAATGGAC 2579
Db 2565 TGGGTGTCNAACCTACTACTCTCTCCAAAGACAAAGACAGCTGTACTTTGTGATGGAC 2624
QY 2580 TACATTCCTGGGGGTGATATGATGAGCCTTAAATTAAGATAGGCACTTTTCCAGAAAGT 2639
Db 2625 TACATCCCTGGTGGGACATGATGAGCCTGCTGATCCGATGGAGCTCTCCCTGAGCAC 2684
QY 2640 CTGGCAGGATTCACATAGACAGAACTTACCTGTCAGTTGAAAGTCTTATATAAATGGGT 2699
Db 2685 CTGGCCCGGTTTACATCGCAGAGCTGACTTTGGCCATTTGAGAGTGTCCACAAAGATGGC 2744
QY 2700 TTATTTATAGAGATATTAACCTGATAATATTTGATTCATGCTGATGCTCATATATAA 2759
Db 2745 TTTATCCAGGACATCAAGCTGTATACATTTGATAGTCTGATGCTGATGCTACATATAA 2804
QY 2760 TTGACTGATTTGGCTCTGCACTGCTCTAGATGACACACAGATTTCTAAGTACTACTAG 2819
Db 2805 CTACAGATTTTGGCCCTCTGCACTGCTGAGTGTGAGTGTGAGTGTGAGTGTGCT 2864
QY 2820 AGTGTGACCATCCAGGCAAGATACATGCTGATTTTCAATTAAGTGGGGGATCCCTCA 2879
Db 2865 AAGGGAGGCGATCTAGACAGGACAGCATGGAGCCGACGCTCTGGGATGATGTGTCT 2924
QY 2880 AGTGTGCGATGTGAGACAGACTGAAGCCATTTAGACGGAGAGCTGCACGCGAGCACCAG 2939
Db 2925 AACTGTGCGTGTGGGACAGAGCTGAGACCCCTAGACAGAGGCGGGAAGCAGCACCAG 2984
QY 2940 CGATGCTAGCACATTTTGTGGTGGGACTTCCCAATTAATGACACTGAAAGTGTGCTA 2999
Db 2985 AGTGTGCTGACATTTCACTGTGGGAGTCCAAACTACATGACACCCGAGGTGCTCTC 3044
QY 3000 CGAACAGGATACACAGATGCTGATGCTGAGTGTGGTGTGCTTATTTTGTGAATG 3059
Db 3045 CGAAAGGGTACACTCAACTCTGTGACTGGTGGAGTGTGGAGTGTATCTCTTCGAGATG 3104
QY 3060 TTGGTGGGACAACTCTTTTCTTGGCACAACACCACTTAGAACACAAATGAAGGTTATC 3119
Db 3105 CTGGTGGGACGCGCCCTTTTGGACCTACTCCACAGAAACCCAGCTGAAGTGTATC 3164
QY 3120 AACTGCAACATCTCTTCCATTCACACCAAGCTAAACTCAGTCCCTGAAAGCTTCTGAT 3179
Db 3165 AACTGGGCAACACGCTCCACATTCAGCCAGGTGAAGTGAAGTCCCTGAGGCCAGGGAC 3224
QY 3180 CTTATTTAATTTGCGGAGGCCGAGATCGTTAGGCAAGAAATGCTGCTGATGAA 3239
Db 3225 CTATCACCAGCTGTGCTGCTGCGAGACACCCGCTGGGCGGAATGGGCGCGATGAC 3284
QY 3240 ATAAAGCTCATCCATTTTTTAAACAAATTTCTCTCAGTGACCTGAGACAGAGCTCT 3299
Db 3285 CTGAAGCCCAACCCCTTCTTCAGGCCATTTGACTTCTCAGTGACATCCGGAAGCAGCCA 3344
QY 3300 GTTCTATACATCTCTTAAATACACACCAACCAAGATACATCAATTTTGTGCTGTGAT 3359
Db 3345 GCGCCCTAGTTCACCACTCAGCCACCCATGACACACCTCGAATTTCCGACCCCGTAGAT 3404
QY 3360 CCTGATAAATTTAGGAGTATGATACGAGGAGAAATGTAATGACACTCTCAATGGA 3419
Db 3405 GAAGAAGCCCTTGGAAACGATGCCAGC---GAAGGTAGCACCAAGGCTTGGGACACACTC 3461

QY 3420 TGTATATAAATGGAAGCATCTGTAACATCTATGAATTTACCTTCCGAAGGTTT 3479
Db 3462 ACCTCCCCCAATTAACAGCATCTGAGCAGCATTTTACGAATTCACCTTCCGAAGGTTT 3521
QY 3480 TTGTATGACAATGGCTACCCATATAATATTCGGAAGCCTATTGAATATGAATACATTAAT 3539
Db 3522 TTGTATGACAATGGCTACCCCTTTTCATGTCGCAAGCCCTTCAGGACGACGAAGTTCACAG 3581
QY 3540 TCACAAGGCTCAGA 3553
Db 3582 GCTGAGAGCTCAGA 3595

RESULT 15
AAZ61160
ID AAZ61160 standard; cDNA; 1961 BP.

XX AAZ61160;
AC AC
XX 30-MAY-2000 (first entry)
XX

DE cDNA SS1771A encoding a partial protein kinase.

XX Kinase activity; molecular weight marker; isoelectric focusing marker;
KW peptide fragmentation control; cellular signal transduction; ss.

XX Homo sapiens.

XX Key Location/Qualifiers
FT CDS 3..1769
FT /tag= a
FT /note= "partial sequence"

XX WO200008180-A2.

XX 17-FEB-2000.

XX 03-AUG-1999; 99WO-US17630.

XX 04-AUG-1998; 98US-0095270.

XX 11-SEP-1998; 98US-0099972.

XX (IMV) IMMUNEX CORP.

PI Virca GD, Bird TA, Anderson DM, Marken JS;

XX WPI; 2000-195584/17.

DR P-PSDB; AAY69162.

XX New human kinase polypeptides and polynucleotides used as molecular weight markers and as controls for peptide fragmentation

PS Claim 1; Page 9; 60pp; English.

XX The present sequence encodes a partial polypeptide which has kinase activity. The kinase polynucleotides can be used to express the polypeptides, and as probes to identify nucleic acids encoding proteins having kinase activity. The kinase polypeptides and fragmented polypeptides are used as molecular weight and isoelectric focusing markers, and as controls for peptide fragmentation. They also have a number of therapeutic uses as kinases play a central role in cellular signal transduction. The polypeptides could also be used to identify binding partner proteins. The polypeptides can also be used as a reagent to identify any proteins that the polypeptide regulates, and proteins with which it might interact. The polypeptides may also be used for preparation of antibodies. The antibodies can be used in assays to detect the presence of the protein, and to purify the protein by immunoaffinity chromatography.

XX Sequence 1961 BP; 504 A; 550 C; 539 G; 368 T; 0 other;

Query Match 18.6%; Score 742.8; DB 21; Length 1961;

Best Local Similarity 66.2%; Pred. No. 1.le-175;

Matches 1122; Conservative 0; Mismatches 557; Indels 15; Gaps 3;

QY 1872 GAAGCTTCCAACTATCAAGGACACACACACACCTCCCAAAACATCTGCT-----GCACC 1926
 Db 21 GGAGGCCAGACGGAGGTGCGCCCTCGCCCTACCCGAAACACCTGCTGCTGCGCAGC 80
 QY 1927 AAAACCCATCTGCTTCCATACGAGTCAATCAGTAAGCCTAGCAAGAGGATCAG---- 1982
 Db 81 AAGTGGAGCAGTACGACCTGGACAGCCTGTGCGAGGATGGAGCAGAGCCTCGGTGG 140
 QY 1983 ---CCAAGCTTCCCAAGCAAGATGAGACTGAAAGAGTATGAAATGTTGATAGTGGG 2039
 Db 141 GCGCCCAAGCGCGGAGGCGGCGGACAGAGCCGCAAGAGCGCAAGGGGAGCAAGGCG 200
 QY 2040 GATAAGAAAAGAAACAGATTACAACTTCCACCTATTACTGTTAGGAAAACAAAGAAAGAT 2099
 Db 201 GGAAGGATAAAAGACAGATTTCAGACCTCTCCCGTCCGCGCAAAACAGAGAGAC 260
 QY 2100 GAAGAGCGAAGGATCTCGTATTCAGAGTTATCTCCTCAAGCAATTTAAATTCCTTTATG 2159
 Db 261 GAAGAGAAGAGAGATCGCATCAAGAGCTACTCGCCATACGCCCTTTAAGTTCTTCATG 320
 QY 2160 GAGCAACATGTAGAAAATGTACTCAAAATCTCATAGCAGCGCTACATCGTAAAAAACAA 2219
 Db 321 GAGCAGCAGTGGAGATGTCAATAACCTACAGCAGAGAGGTTAAACGGAGGCTGCAG 380
 QY 2220 TTAGAGAAATGAATGATCGGGTTCGATTATCTCAAGATGCCCAGGATCAATGAGAAAG 2279
 Db 381 CTGGAGCAAGAAATGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATCGGAG 440
 QY 2280 ATGCTTTCGCAAAAGATCTAATTTACATCGCTTTAAAGGGCTAAATGGACAAAGTCT 2339
 Db 441 ATCTCTTACCAAGAAAGATCTAATTTACACAGGTTAAAGAGGGCCAAAGATGGACAAAGTCT 500
 QY 2340 ATGTTTGTGAAGATAAAGACACTAGGATAGGACATTTGTTGAAGTCTGTCTAGCAAGA 2399
 Db 501 ATGTTTGTCAAGTCAAAACCTGGGATCGGTGCTTTGGAGAAGTGTGCTGCTTGT 560
 QY 2400 AAGTAGATATAAGGCTTTGTATGCAACAAACCTCTTCGAAAGAAAGATGTTCTTCTT 2459
 Db 561 AAGTGGACACTACGCCCTGTACGCCATGAAGACCTTAAGGAAAGAGATGCTCTGAAC 620
 QY 2460 GAAATCAAGTCTGCTATGTTAAGCTCAGAGAGATATCTGGCTGAAGCTGACAAATGAA 2519
 Db 621 CGGAATCAGGTGGGCCCGCTCAAGCCGAGAGGGGACATCTTGGCCGAGGAGCAATGAG 680
 QY 2520 TGGTGTGCTGTATATTTATTTCAATCCAAAGTAAGCAATTTATCTTTGTTAATGGAC 2579
 Db 681 TGGTGTGCTCAAACTCTACTCTCTTCCAAAGCAAGACAGCCTGTACTTTGTGATGGAC 740
 QY 2580 TACATTCCTGGGGTGATATGATGAGCCTATTAATTAGAATGGGCATCTTTCCAGAAAGT 2639
 Db 741 TACATCCCTGGTGGGGATGATGAGCCTGCTGATCCGGATGGAGTCTTCCCTGAGCAC 800
 QY 2640 CTGGCAGCATTTACATAGCAGAACTTACCTGTGCAAGTTGAAGTGTCAATAAATGGGT 2699
 Db 801 CTGGCCCGGTTCTACATCGCAGAGTGAATTTGGCCATTTGAGAGTGTCCACAAAGATGGC 860
 QY 2700 TTTATTCATAGAGATATTAACCTGTATTAATTTTGTGATCGTGTGTTGATTAATAA 2759
 Db 861 TTCATCCACGAGACATCAACCTGTATTAACATTTTGTAGATCTGGATGGTACATTA 920
 QY 2760 TTGACTGACTTTGGCTCTGCACTGGGTTTCAAGTGGACACAGATTTCTAAGTACTATCAG 2819
 Db 921 CTCACAGATTTCCGGCTCTGCACTGGGTTTCAAGTGGACTACAATTTCAATAATTACCAG 980
 QY 2820 AGTGTGACCATCCAGCAGATGAGATGATTTTCAATGATGATGGGGATCCCTCA 2879
 Db 981 AAAGGAGCCATGTACAGAGAGAGATGAGAGCCATTAGAGGGGAGAGCTGACCGCAGCAG 2939
 QY 2880 AGCTGTGATGTGGAGACAGACTGAAGCCATTAGAGGGGAGAGCTGACCGCAGCAGCAG 2939

Db 1041 AACTCTCGTGTGGGACAGGCTGAAGACCTTAGACAGAGCGCGGAAAGCAGCAGCAG 1100
 QY 2940 CGATCTCTAGCACATTTCTTTGGTGGGACTCCCAATTTATATGACCTGAAGTGTGCTA 2999
 Db 1101 AGTGCCTGGACATCTACTGGTGGGACTCCAACTACATCGCACCCGAGGTGCTCTC 1160
 QY 3000 GGAACAGGATACACACAGTGTGTGATTGGTGGAGTGTGGTGTATTTCTTTTGAATG 3059
 Db 1161 CGAAAGGATACACTCAACTCTGTGACTGGTGGAGTGTGGAGTGTCTCTTCGAGATG 1220
 QY 3060 TGTGGGACAAACCTCTCTTTCTGGCACAACACCACTTAGAAACACAATGAAGTTATC 3119
 Db 1221 CTGTGGGGGAGCGCCCTTTTGGCACTTCTCCACAGAAACCCAGCTGAAGGTGATC 1280
 QY 3120 AACTGGCAACATCTCTTTCACATTCACCAACCAAGCTAACTCAGTCCCTGAAGCTTCTGAT 3179
 Db 1281 AACTGGGAGACACGCTCCACATTCACGCCCCAGGTGAAGCTGAGCCCTGAGCCAGGAC 1340
 QY 3180 CTTATTATTAACCTTTGCCGAGGACCCGAAGATCGCTTTAGGCAAGAATGTTGCTGATGAA 3239
 Db 1341 CTCATCACCAAGCTGTGCTGCTCCGACAGACACCGCTGGGGCGGAATGGGGCGATGAC 1400
 QY 3240 ATAAAGCTCATCTCTTTTAAACAAATTCAGTCTCCAGTGGACCTGAGACAGCAGTCT 3299
 Db 1401 CTGAAGGCCCCCCTCTCTTCAGGCGCATTTGACTTCTCCAGTGACATCCGGAAGCATCCA 1460
 QY 3300 GCTTCATACATTCCTAAATCACACACCAACACATACATCAAAATTTTGTCTGTTGAT 3359
 Db 1461 GCGCCCTACGTTCCACCATCAGCCACCCCATGGACACCTGGAATTTGACCCCGTAGAT 1520
 QY 3360 CCGTGAATAATTTGAGTGTGATGAACGAGGAAGAAATGTAATGACACTCTCAATGGA 3419
 Db 1521 GAAGAAAGCCCTTGGAGGATGCCAGC---GAAGTAGCACCAAGGCTGGGACACACTC 1577
 QY 3420 TGGTATAAATATGAAAGCATCTTGAACATGCTATGATTTTACCTTCCGAGGTTT 3479
 Db 1578 ACCTGCGCCCAATCAAGCATCTCTGAGCAGCAGCTTTTACGAATTTACCTTCCGAGGTTT 1637
 QY 3480 TTTGATGACATGGCTACCCCATATAATTTATCCGAAGCCTATTGAATATGAATACATTAAT 3539
 Db 1638 TTTGATGACATGGCTACCCCTTTTCGATGCCCAAGGCTTCAGGAGCAGAGCTTCACAG 1697
 QY 3540 TCACAAGGCTCAGA 3553
 Db 1698 GCTGAGAGCTCAGA 1711

Search completed: January 16, 2003, 10:32:01
 Job time : 562.163 secs

Result No.	Query			DB	ID	Description
	Score	Match	Length			
1	3984	100	0	3984	4	US-09-442-100-3
2	2483	62	3	3213	4	US-09-442-100-5
3	760	19	1	3155	4	US-09-442-100-7
4	742.8	18	6	1961	4	US-09-509-902A-15
5	675.8	17	0	1498	4	US-09-509-902A-6
6	609.8	15	3	638	4	US-09-328-111-26
7	469	11	8	5720	4	US-09-442-100-1
8	452.8	11	4	678	4	US-09-328-111-66
9	269.4	6	8	1935	2	US-08-878-989-11
10	269.4	6	8	1935	4	US-09-272-796-11
11	254.6	6	4	3018	2	US-08-860-150-6
12	254.6	6	4	3018	3	US-09-338-132-6
13	173.2	4	3	2160	4	US-09-588-256-1
14	172.4	4	3	2101	2	US-08-860-150-1
15	172.4	4	3	2101	3	US-09-338-132-1
16	136.2	3	4	4739	2	US-08-685-871-1
17	122.8	3	1	2706	2	US-08-630-822A-61
18	122.8	3	1	2706	2	US-09-005-069-61
19	122.8	3	1	2706	4	US-09-171-156A-20
20	121.8	3	1	5053	2	US-08-685-576-2
21	117	2	9	4363	2	US-08-685-576-5
22	111	2	8	2311	2	US-08-712-709-6
23	111	2	8	2311	3	US-09-111-444-6
24	111	2	8	2311	4	US-09-541-228-6
25	111	2	8	2370	4	US-09-031-295-1
26	104	2	6	1599	4	US-09-256-465-1
27	104	2	6	1599	4	US-09-167-322-3
28	104	2	6	1599	4	US-09-167-322-3
29	104	2	6	1599	4	US-09-167-322-3
30	104	2	6	1599	4	US-09-167-322-3
31	104	2	6	1599	4	US-09-167-322-3
32	104	2	6	1599	4	US-09-167-322-3
33	104	2	6	1599	4	US-09-167-322-3
34	104	2	6	1599	4	US-09-167-322-3
35	104	2	6	1599	4	US-09-167-322-3
36	104	2	6	1599	4	US-09-167-322-3
37	104	2	6	1599	4	US-09-167-322-3
38	104	2	6	1599	4	US-09-167-322-3
39	104	2	6	1599	4	US-09-167-322-3
40	104	2	6	1599	4	US-09-167-322-3
41	104	2	6	1599	4	US-09-167-322-3
42	104	2	6	1599	4	US-09-167-322-3
43	104	2	6	1599	4	US-09-167-322-3
44	104	2	6	1599	4	US-09-167-322-3
45	104	2	6	1599	4	US-09-167-322-3
46	104	2	6	1599	4	US-09-167-322-3
47	104	2	6	1599	4	US-09-167-322-3
48	104	2	6	1599	4	US-09-167-322-3
49	104	2	6	1599	4	US-09-167-322-3
50	104	2	6	1599	4	US-09-167-322-3
51	104	2	6	1599	4	US-09-167-322-3
52	104	2	6	1599	4	US-09-167-322-3
53	104	2	6	1599	4	US-09-167-322-3
54	104	2	6	1599	4	US-09-167-322-3
55	104	2	6	1599	4	US-09-167-322-3
56	104	2	6	1599	4	US-09-167-322-3
57	104	2	6	1599	4	US-09-167-322-3
58	104	2	6	1599	4	US-09-167-322-3
59	104	2	6	1599	4	US-09-167-322-3
60	104	2	6	1599	4	US-09-167-3

US-09-442-100-3

Query Match.	100.0%;	Score 3984;	DB 4;	Length 3984;
Best Local Similarity	100.0%;	Pred. No. 0;		
Matches 3984;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
QY	1	ACCTTTGGTGTCTGGACGACACTCTGGCGCCCTCAGCTCGCCCTCAGCGCCGTGGCC	60	
DB	1	ACCTTTGGTGTCTGGACGACACTCTGGCGCCCTCAGCTCGCCCTCAGCGCCGTGGCC	60	
QY	61	GCTGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATATTGTAAGAATTTTA	120	
DB	61	GCTGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTTATTTATATTGTAAGAATTTTA	120	
QY	121	ACAGTCTGGGGACTTCTCTTGAAGATCATTTTCACATTTTCTCAGAGAAGCTCTGGA	180	
DB	121	ACAGTCTGGGGACTTCTCTTGAAGATCATTTTCACATTTTCTCAGAGAAGCTCTGGA	180	
QY	181	TCTATCAAAATAAGAAAGTCTCTTCTGTGGGTACATATATAGATGTTTTCATGAAGAGGA	240	
DB	181	TCTATCAAAATAAGAAAGTCTCTTCTGTGGGTACATATATAGATGTTTTCATGAAGAGGA	240	
QY	241	GTGAAAAGCCAGAAGGATATAGACAAATGAGGCTTAAGACCTTTCTGCCAGTAACATATA	300	
DB	241	GTGAAAAGCCAGAAGGATATAGACAAATGAGGCTTAAGACCTTTCTGCCAGTAACATATA	300	
QY	301	CTGTCACTAGCGGCAAAATGTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC	360	
DB	301	CTGTCACTAGCGGCAAAATGTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAAC	360	
QY	361	CATCTGATGCTGTAAAGGCTGAGCATAACTAGTAAATGTCAACCGAAGATCCTCGAC	420	
DB	361	CATCTGATGCTGTAAAGGCTGAGCATAACTAGTAAATGTCAACCGAAGATCCTCGAC	420	
QY	421	AAGTCAGAAATCCACCANAATTTGGGAGCATCATAAAGCCTTCGAGGAATTCGAAACT	480	
DB	421	AAGTCAGAAATCCACCANAATTTGGGAGCATCATAAAGCCTTCGAGGAATTCGAAACT	480	
QY	481	CTCTGCTTCCATTTGCAAAATGAAACAAATTTCTCTCGGAGTACCTCAGAACTTAATCCAC	540	
DB	481	CTCTGCTTCCATTTGCAAAATGAAACAAATTTCTCTCGGAGTACCTCAGAACTTAATCCAC	540	
QY	541	AAATGCTTCAAGACTTCCAGCTGCTGGATTTGATAGGATATGGTATATCAAGCTCTTC	600	
DB	541	AAATGCTTCAAGACTTCCAGCTGCTGGATTTGATAGGATATGGTATATCAAGCTCTTC	600	
QY	601	AGAAACTCAACACAGAAATGATAGACGACCAATTTGAATTCATTAGTAAATGAGTTACC	660	
DB	601	AGAAACTCAACACAGAAATGATAGACGACCAATTTGAATTCATTAGTAAATGAGTTACC	660	
QY	661	AAGATCTTCAGCAGAGCAGATGGCTGCAGCAGCTGCCAGACCTATTAAATGCCAGCATGA	720	
DB	661	AAGATCTTCAGCAGAGCAGATGGCTGCAGCAGCTGCCAGACCTATTAAATGCCAGCATGA	720	
QY	721	AACCAGGAATGTGCAGCAATCAGTTAACGCCAACAGAGCTGGAAGGTTCTTAAAGAAAT	780	
DB	721	AACCAGGAATGTGCAGCAATCAGTTAACGCCAACAGAGCTGGAAGGTTCTTAAAGAAAT	780	
QY	781	CCTTAGTCTTCAGAGCATGGCCGCCCTTAGGAGAAAGTGTGGCCTATCATTTCTCAGA	840	
DB	781	CCTTAGTCTTCAGAGCATGGCCGCCCTTAGGAGAAAGTGTGGCCTATCATTTCTCAGA	840	
QY	841	GTCCCAACTCACAGACAGATGTAGGAAGACTTTGTCTGGATCTGGTATATCAGCATTTG	900	
DB	841	GTCCCAACTCACAGACAGATGTAGGAAGACTTTGTCTGGATCTGGTATATCAGCATTTG	900	
QY	901	TTCAAGCTCACTTAGCAACGGACAGAGTGAACCCCCCACCACCACTCAAGTAAGGA	960	
DB	901	TTCAAGCTCACTTAGCAACGGACAGAGTGAACCCCCCACCACCACTCAAGTAAGGA	960	
QY	961	GTCTTACTCTCCACCACCTCCCAAGAGGCCAGACTCTCCCTTCCAAGAGGTACAACTCCAC	1020	
DB	961	GTCTTACTCTCCACCACCTCCCAAGAGGCCAGACTCTCCCTTCCAAGAGGTACAACTCCAC	1020	

Qy	1021	CTCCCCCTTCATGGGAACCAAACTCTCAACAAAGCGGTATTTCTGGAAACAATGAATACG	1080
Db	1021	CTCCCCCTTCATGGGAACCAAACTCTCAACAAAGCGGTATTTCTGGAAACAATGAATACG	1080
Qy	1081	TAATCTCCCGAATCTCTCTGTGCCACCTGGGCGCATGCCAAGAGGGCTATCTCTCCACCAC	1140
Db	1081	TAATCTCCCGAATCTCTCTGTGCCACCTGGGCGCATGCCAAGAGGGCTATCTCTCCACCAC	1140
Qy	1141	CTCTCAACACTTCCCCCATGAATCTCTTAATCAAGGACAGAGAGCATTAGTTCTGTTC	1200
Db	1141	CTCTCAACACTTCCCCCATGAATCTCTTAATCAAGGACAGAGAGCATTAGTTCTGTTC	1200
Qy	1201	CTGTGGCAGACAACCAATCATGCGAGATTCATAGCAAAATTTAACTTCCATCAGGGA	1260
Db	1201	CTGTGGCAGACAACCAATCATGCGAGATTCATAGCAAAATTTAACTTCCATCAGGGA	1260
Qy	1261	GACCTGGAATCGAATGGTACTTGGACAACATGATTTTCATGATACACCAAAATGTTCTCC	1320
Db	1261	GACCTGGAATCGAATGGTACTTGGACAACATGATTTTCATGATACACCAAAATGTTCTCC	1320
Qy	1321	CTGCTGGCACTGTGAATCGGCAGCCACACCTCCATATCTCTTGACAGCAGCTAATGGAC	1380
Db	1321	CTGCTGGCACTGTGAATCGGCAGCCACACCTCCATATCTCTTGACAGCAGCTAATGGAC	1380
Qy	1381	AAAGCCCTTCCTGCTTTTACAAAAGGGGATCGCTGCTCTTCCTGCTCATATACAAATGGA	1440
Db	1381	AAAGCCCTTCCTGCTTTTACAAAAGGGGATCGCTGCTCTTCCTGCTCATATACAAATGGA	1440
Qy	1441	GTATTCCTCAGTCTATGATGGTGCCAAACAGAAATAGTCATAACATGGAATATATAACA	1500
Db	1441	GTATTCCTCAGTCTATGATGGTGCCAAACAGAAATAGTCATAACATGGAATATATAACA	1500
Qy	1501	TTAGTGTACCTGGACTGCAAAACAATTTGGCCTCAGTCATCTTCTGCTCCAGCCCGAGTCAT	1560
Db	1501	TTAGTGTACCTGGACTGCAAAACAATTTGGCCTCAGTCATCTTCTGCTCCAGCCCGAGTCAT	1560
Qy	1561	CCCCGAGCAGTGGGCATGAAATCCCTACATGCAACCTTAAACATACCAGTCCAGGTCAAAT	1620
Db	1561	CCCCGAGCAGTGGGCATGAAATCCCTACATGCAACCTTAAACATACCAGTCCAGGTCAAAT	1620
Qy	1621	CTTTTAAATACCACTTAGGAATAGAGCAAGTCACTCTGCTTAATTCACGCTTCGCTA	1680
Db	1621	CTTTTAAATACCACTTAGGAATAGAGCAAGTCACTCTGCTTAATTCACGCTTCGCTA	1680
Qy	1681	CAACAGTCACTGCAATTTACACAGCTCCTTATTCACAGCCCTGTGAAAGATATGCGTGTAT	1740
Db	1681	CAACAGTCACTGCAATTTACACAGCTCCTTATTCACAGCCCTGTGAAAGATATGCGTGTAT	1740
Qy	1741	TAAACACAGAGCTACAGACTGCTTTTAGCACCTTACACACCCCTTCCTTGGATACCAACGCCAA	1800
Db	1741	TAAACACAGAGCTACAGACTGCTTTTAGCACCTTACACACCCCTTCCTTGGATACCAACGCCAA	1800
Qy	1801	TTCAAACTGTTCAACCAGCTCTTTTCTGAGGGAACCGCTTCAAAATGTCAGTGTGATGC	1860
Db	1801	TTCAAACTGTTCAACCAGCTCTTTTCTGAGGGAACCGCTTCAAAATGTCAGTGTGATGC	1860
Qy	1861	CACCTGTGCTGAAGCTCCAAACTATCAAGGACACCAACCCCTACCCAAAACATCTGC	1920
Db	1861	CACCTGTGCTGAAGCTCCAAACTATCAAGGACACCAACCCCTACCCAAAACATCTGC	1920
Qy	1921	TGCACAAAACCCATCTGTTCTCCATACGAGTCAATCAGTTAAGCTTAGCAAGAGGATC	1980
Db	1921	TGCACAAAACCCATCTGTTCTCCATACGAGTCAATCAGTTAAGCTTAGCAAGAGGATC	1980
Qy	1981	AGCAAGCTTGCCCAAGGAAGATGAGTGAAGAGGTTATGAAAATGTTGATAGTGGGG	2040
Db	1981	AGCAAGCTTGCCCAAGGAAGATGAGTGAAGAGGTTATGAAAATGTTGATAGTGGGG	2040
Qy	2041	ATAAGAAAGAAACAGATTACAACCTTCACTTATTCTGTTAGGAAAAACAAGAAAGATG	2100
Db	2041	ATAAGAAAGAAACAGATTACAACCTTCACTTATTCTGTTAGGAAAAACAAGAAAGATG	2100

QY 2101 AAGAGCGAAGGGAATCTCGTATTCAAAGTTATTCTCTCAAGCATTTAAATCTTTATGG 2160
 DB 2101 AAGAGCGAAGGGAATCTCGTATTCAAAGTTATTCTCTCAAGCATTTAAATCTTTATGG 2160
 QY 2161 AGCAACATGTAGAAATGTACTCAATCTCATCAGAGCGTCTACATCGTAAAAACAAT 2220
 DB 2161 AGCAACATGTAGAAATGTACTCAATCTCATCAGAGCGTCTACATCGTAAAAACAAT 2220
 QY 2221 TAGAGAAATGAATGATCGGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGA 2280
 DB 2221 TAGAGAAATGAATGATCGGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGA 2280
 QY 2281 TGCTTTGCCAAAAGAAATCTAATTACATCGTCTTAAAGGGCTTAAATGGACAAGTCTA 2340
 DB 2281 TGCTTTGCCAAAAGAAATCTAATTACATCGTCTTAAAGGGCTTAAATGGACAAGTCTA 2340
 QY 2341 TGTTTGTGAAGATTAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGA 2400
 DB 2341 TGTTTGTGAAGATTAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGA 2400
 QY 2401 AAGTAGATCTAAGGCTTGTATGCAACAAAACTCTCCGAAAGAAAGATGTTCTTCTC 2460
 DB 2401 AAGTAGATCTAAGGCTTGTATGCAACAAAACTCTCCGAAAGAAAGATGTTCTTCTC 2460
 QY 2461 GAAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCTCGGCTGAAGCTCACAATGAAT 2520
 DB 2461 GAAATCAAGTCGCTCATGTTAAGGCTGAGAGATATCTCGGCTGAAGCTCACAATGAAT 2520
 QY 2521 GGGTAGTCTGTATATTAATCAATCAAGATTAAGGCAATTAATCTTTGTAATGAGCT 2580
 DB 2521 GGGTAGTCTGTATATTAATCAATCAAGATTAAGGCAATTAATCTTTGTAATGAGCT 2580
 QY 2581 ACATTCTGGGGTGATGATGAGCCTATTAAATGAATGGGCACTTTCCAGAAAGTC 2640
 DB 2581 ACATTCTGGGGTGATGATGAGCCTATTAAATGAATGGGCACTTTCCAGAAAGTC 2640
 QY 2641 TGCCAGCATCTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTT 2700
 DB 2641 TGCCAGCATCTACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTT 2700
 QY 2701 TTATTCATAGAGATTAATCACTGATAATTTTGTGATCGTGATGGTCATATTAAT 2760
 DB 2701 TTATTCATAGAGATTAATCACTGATAATTTTGTGATCGTGATGGTCATATTAAT 2760
 QY 2761 TGACTGACTTTGGCCCTGTCAGCTGGCTTCAGATGGACACACGATTTCTAAGTACTATCAGA 2820
 DB 2761 TGACTGACTTTGGCCCTGTCAGCTGGCTTCAGATGGACACACGATTTCTAAGTACTATCAGA 2820
 QY 2821 GTGGTGACCATCCACGGCAAGATAGCATGGATTTCAGTAATGAATGGGGGATCCCTCAA 2880
 DB 2821 GTGGTGACCATCCACGGCAAGATAGCATGGATTTCAGTAATGAATGGGGGATCCCTCAA 2880
 QY 2881 GCTGTGATGTGAGACAGACTGAAGCCATTAGAGGGAGAGCTGCACGCCAGCAGCAGC 2940
 DB 2881 GCTGTGATGTGAGACAGACTGAAGCCATTAGAGGGAGAGCTGCACGCCAGCAGCAGC 2940
 QY 2941 GATGTCTAGCACATCTTTGGTGGGACTCCCAATTAATTTGCACCTGAAGTGTGCTAC 3000
 DB 2941 GATGTCTAGCACATCTTTGGTGGGACTCCCAATTAATTTGCACCTGAAGTGTGCTAC 3000
 QY 3001 GAACAGGATACACAGATTTGTGATTTGGTGGAGTGTGGTGTATCTTTTGAATGT 3060
 DB 3001 GAACAGGATACACAGATTTGTGATTTGGTGGAGTGTGGTGTATCTTTTGAATGT 3060
 QY 3061 TGTTGGGACACCTCTTTCTGGCACAACACCATTTAGAAACACAAATCAAGGTTATCA 3120
 DB 3061 TGTTGGGACACCTCTTTCTGGCACAACACCATTTAGAAACACAAATCAAGGTTATCA 3120
 QY 3121 ACTGGCAACATCTTTACATTTCCACCAAGCTAACTCACTGCTGAAGCTTCTGATC 3180
 DB 3121 ACTGGCAACATCTTTACATTTCCACCAAGCTAACTCACTGCTGAAGCTTCTGATC 3180
 QY 3181 TTATTATTAACTTTGCCGAGGACCGAAGATCGCTTAGCAAGAATGGTGTGATGAAA 3240

DB 3181 TTATTATTAACTTTGCCGAGGACCCCAAGATCGTTAGCAAGAATGGTGTGATGAAA 3240
 QY 3241 TAAAGCTCATCTCAATTTTAAACAATTTGACTTCTCCAGTGACCTGAGACAGAGTCTG 3300
 DB 3241 TAAAGCTCATCTCAATTTTAAACAATTTGACTTCTCCAGTGACCTGAGACAGAGTCTG 3300
 QY 3301 CTTCATACATTTCTTAAATCACACACCACAGATACATCAAAATTTTGTCTTCTGATC 3360
 DB 3301 CTTCATACATTTCTTAAATCACACACCACAGATACATCAAAATTTTGTCTTCTGATC 3360
 QY 3361 CTGATAAATTTATGAGTGATGATAACGAGGAAGAAATGTAAATGACACTCTCAATGGAT 3420
 DB 3361 CTGATAAATTTATGAGTGATGATAACGAGGAAGAAATGTAAATGACACTCTCAATGGAT 3420
 QY 3421 GGTATAAATTTGAAAGCATCTCGAATCATCTTCTATGAATTTACCTCCGAAGGTTTT 3480
 DB 3421 GGTATAAATTTGAAAGCATCTCGAATCATCTTCTATGAATTTACCTCCGAAGGTTTT 3480
 QY 3481 TTGATGACATGCTACCCATATAATTTATCCGAAGCCTATTGAATATGAATACATTAAT 3540
 DB 3481 TTGATGACATGCTACCCATATAATTTATCCGAAGCCTATTGAATATGAATACATTAAT 3540
 QY 3541 CACAAGGCTCAGAGCAGCAGTCCGATGAAGATGATCAAAAACACAGGCTCAGAGATTTAAA 3600
 DB 3541 CACAAGGCTCAGAGCAGCAGTCCGATGAAGATGATCAAAAACACAGGCTCAGAGATTTAAA 3600
 QY 3601 ATCCGATCTAGTATATGTTTAAACACACTAGTAAATTAATGATGAGGATTTGTAAG 3660
 DB 3601 ATCCGATCTAGTATATGTTTAAACACACTAGTAAATTAATGATGAGGATTTGTAAG 3660
 QY 3661 GGCTGAAATGCGAGGTTTGGAGGTTCTGAGAGTAAATTTATGCAAAATATGACAGC 3720
 DB 3661 GGCTGAAATGCGAGGTTTGGAGGTTCTGAGAGTAAATTTATGCAAAATATGACAGC 3720
 QY 3721 TATATATGCTGCTGTGTACAAATTTTATTTTCTTAAATTTATGCGAAATCCTTTTAA 3780
 DB 3721 TATATATGCTGCTGTGTACAAATTTTATTTTCTTAAATTTATGCGAAATCCTTTTAA 3780
 QY 3781 AATGTTAAATTTATCCAGCGTTTAAATCAGTATTTAGAAAAAATGTTTAAAGGAAG 3840
 DB 3781 AATGTTAAATTTATCCAGCGTTTAAATCAGTATTTAGAAAAAATGTTTAAAGGAAG 3840
 QY 3841 TAAATATGAACCTGAATATATAGTCTGAGTCTTGGTACTTAAAGTACTTAAATAGTAG 3900
 DB 3841 TAAATATGAACCTGAATATATAGTCTGAGTCTTGGTACTTAAAGTACTTAAATAGTAG 3900
 QY 3901 TGCTTTGTTTAAAGGAGAAACCTGATCTATTTGATATATGCTAAATTTTAAAA 3960
 DB 3901 TGCTTTGTTTAAAGGAGAAACCTGATCTATTTGATATATGCTAAATTTTAAAA 3960
 QY 3961 TACAAGAGTTTTTCAAAATTTTTT 3984
 DB 3961 TACAAGAGTTTTTCAAAATTTTTT 3984

RESULT-2

US-09-442-100-5

Sequence 5, Application US/09442100

Patent No. 6359193

GENERAL INFORMATION:

APPLICANT: Xu, Tian

APPLICANT: Wang, Welyi

APPLICANT: Zhang, Sheng

APPLICANT: Yu, Wan

TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0; Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442.100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 5:
SEQUENCE CHARACTERISTICS:
LENGTH: 3213 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2889
US-09-442-100-5

Query Match 62.3%; Score 2483; DB 4; Length 3213;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;

QY	732	GTGCGAATCATGTTAAACCGCAAGAGCTGGAAGGTTCTAAAGATCTTACTTCT	791
DB	1	GTGCAACATTCAATTAACCGCAAGAGCTGGAAGGTTCTAAAGAGTCTTACTTCT	60
QY	792	CAGAGCATGGCCGCTAGGAGAAAGTGGGCTATCATCTGAGAGTCCCAACTCA	851
DB	61	CAGAGACAGGGCCATCTCTAGAGAAATGTTGTTATCTGTAAGCCCAACTCA	120
QY	852	CAGACAGATGTAGGAAGACCTTTGTCTGGATCTGTGTATATACAGCATTTG	911
DB	121	CAGCGGATGTAGGAAGACCTCTGTCTGGATCCGGCATTTGCAGCATTTGCT	180
QY	912	CCTAGCAAGGACAGAGTGAACCCGCCACCAACCTCAAGTAGGAGTGTACTCT	971
DB	181	CCAAGCAATGGACAGAGTGAACCCGCCACCAACCTCAAGTAGGAGTGTACTCT	240
QY	972	CCACCACTCCAAAGAGCCAGACTCCCGCTCCCAAGAGTACAACTCCACCTCC	1031
DB	241	CCACCACTCCGAGAGCCAGACCCACCTCCCGGAGGACCACTCCCGCTCCCG	300
QY	1032	TGGGAACCAACTCTCAACAAAGCGCTATTCTGGAACATGGAATACGTACTCC	1091
DB	301	TGGGAACCAAGCTCTCAGACAAAGCGCTATCTGGAACATGGAATACGTACTCC	360
QY	1092	ATCTCTCTCTCCACTGGGCTAGGAGGCTATCTCCACCACTCTCAACACT	1151
DB	361	ATCTCTCTCTCTCCACTGGGCTAGGAGGCTATCTCCACCACTCTCAACACT	420
QY	1152	TCGCCCATGAATCTCTTAATCAAGACAGAGAGGATTTAGTTCTGTCTGTGCGA	1211
DB	421	TCGCCCATGAATCTCTTAATCAAGACAGAGAGGATTTAGTTCTGTCTGTGCGA	480
QY	1212	CAACCAATCATGACAGAGTTCTAGCAATTTAACTTCCATCAGGAGACCTGGAATG	1271

481 CAACCCATCATCATGAGAGTACTAGCAAAATTTAACTTTACACGAGGCGACCTGGAGTT 540
1272 CAGAAATGGTACTGGCAAAAGTGAATTCATGATACACCAAAATGTTGCTCCTGCTGCACT 1331
541 CAGAAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 597
1332 GTGAATCGGAGGACCAACCTCCATATCTCTGACAGAGCTAATGGACAAAGCCCTTCT 1391
598 GTGACTCGGAGGACCAACCTCCATATCTCTGACCCAGCTAATGGACAAAGCCCTTCT 657
1392 GCTTTACAAACAGGGGATCTGCTGCTCTGCTATATACAAATGGAAGTATTCCTCAG 1451
658 GCTTTACAAACAGGGGCTTCTGCTGCTCCACCATCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 717
1452 TCTATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1511
718 TCGATGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 777
1512 GGACTGCAAAACAAATTTGGCTCAGTCTATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1571
778 GGACTGCAAAACAGCTTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
1572 GGGCATGAATTCCTTACATGGCAACCTAACATACCACTGAGGTCATAATCTTTTAAATAAC 1631
838 GGGCATGAATTCCTTACATGGCAACCTAACATACCACTGAGGTCATAATCTTTTAAATAAC 897
1632 CCATTAGAAATAGACAAAGTCACTCTGCTAAATTTCTCAGCCTTCTGCTACAAAGTCACT 1691
898 CCATTAGAAATAGACAAAGTCACTCTGCTAAATTTCTCAGCCTTCTGCTACAAAGTCACT 957
1692 GCAATTACACAGCTCTTATTCACAGCTGTGAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1751
958 GCCATCACACCGCTCTTATTCACAGCTGCTGAAAGTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1017
1752 CTACAGAGTCTTTAGCACCTACACACCTCTTTGGATACCAACCACTTCAAACTGTT 1811
1018 CTGACAGTCTTTAGCCCCCAACCTATCTTTGGATGCCACAGCCTTACAGCTGTT 1077
1812 CAACCCAGTCTTTCTGAGGAAACCGCTTCAAAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1871
1078 CAGCTACCCCTTTCTGAGGTAAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1137
1872 GAAGTCCAAACTATCAAGGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1931
1138 GAAGTCCAAACTATCAAGGACCAACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1197
1932 CCATCTGTTCTCCATGAGTCAATCAGTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1991
1198 CCATCTGCTCCATATGAGTCAAGTAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1257
1992 CCCAAGGAGATGAGAGTGAAGAGTATGAAATGTTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2051
1258 CCCAAGGAGATGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGTGAAGAGT 1317
2052 AAACAGATTACAACTTCACTTACTTCTGTAAGAAACCAAGAAAGATGAGAGGCAAGG 2111
1318 AAACAGATTACAACTTCACTTACTTCTGTAAGAAACCAAGAAAGATGAGAGGCAAGG 1377
2112 GAATCTCTGTTTCAAAAGTATTTCTCTCAAGCATTTAAATTTTATGAGCAACATGTA 2171
1378 GAGTCTCGGATTCAGAGTTACTTCCCAACAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1437
2172 GAAATGTACTCAATCTCATCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2231
1438 GAGAACGCTCTGAAGTCTCATCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1497
2232 ATGATCGGGTGGATTTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAA 2291
1498 ATGATCGGGTGGATTTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAA 1557
2292 AAAGATCTTAATTTACATCTGCTTTAAAGGGCTTAAATGAGAAAGATGCTTTGTTGAAG 2351
1558 AAAGATCTTAATTTACATCTGCTTTAAAGGGCTTAAATGAGAAAGATGCTTTGTTGAAG 1617

Qy	3432	GGAAAGCATCCTGAACATGATCTTATGAATTTACCTTCCGGAAGGTTTTTTGATGACAAT	3491
Db	2698	GGGAAGCACCCCGAGCAGCGCTTCTATGAGTTTCACGCTTCGAGGGTTTTTTGATGACAAT	2757
Qy	3492	GGCTACCCTATTAATTTATCCGAAGCCTATTGAATATGAATACATTAATTACAAAGGCTCA	3551
Db	2758	GGCTACCCTATTAATTTATCCAAAGCCTATTGAGTATGAATACATTCATTCACAGGGCTCA	2817
Qy	3552	GAGCAGCGTCGGATGAAGATGATCAAAACACAGGCTCAGAGATTAAAAATCGCGATCTA	3611
Db	2818	GAACACAGCTCTGATCAAGATGATCAACACACAAAGCTCCGATGGAAACAACCGAGATCTA	2877
Qy	3612	GTATATGTTTTAACACACTAGTAAATAAATGTAATCAGAGTTGTAAAGGCGCCTCAATG	3671
Db	2878	GTGTATGTTTTAATAAACTAGGAGATCATTTGTAA-----GAATTTGCAAGGCGCCTGAAGTG	2933
Qy	3672	CGAGGTGTTTTTGAGGTTCTGAGAGTAAATTTATGCAAAATATCAGACAGCTATATATGCT	3731
Db	2934	CAGGGGTTTTTGAAGTTTTTGAG---AAAATTTATGCAAAATGTCACAGAG-----TTTCTGT	2985
Qy	3732	GCCTCTGTGTACATAATTTTTATTTTCCCTAAATTTATCGGGAATCCCTTTTAAATGTTAAATTT	3791
Db	2986	GCCTCTGTGTACATAATTTTTATTTTCCCTAAAGTTATGGGAAATGTTTTTAAATGTTAAATTT	3045
Qy	3792	ATTCCAGCCGTTTTAAATCAGTATTTTAGAAAAAAATTTGTTAAGGAAAGTAAATTTATGAA	3851
Db	3046	ATTCCACCCCTTTTAAATTCAGTAAATTTAGAAAAAAATGTTTAAAGGAAGTAATTTATGAA	3105
Qy	3852	CTGAATATATAGTCAGTCTCTGGTACTTTAAAGTACTTTAAATTAAGTAG	3900
Db	3106	CTGAGTATATAGTCAATCTCTGGTACTTTAAAGTACTTTAAAGGAAGAG	3154

US-09-44

; sequence 1, APPLICATION US/094442100
: Patent No. 6250102

; GENERAL INFORMATION

; APPLICANT: Tao, W

APPLICANT: Zhang, shen

APPLICANT: IU, WASH	TITLE OF INVENTION:

TITLE OF INVENTION:

CORRESPONDENCE ADDRESS:

STREET: 1155 Avenue of the

CITY: NEW YORK
STATE: NEW YORK

; COUNTRY: USA

; COMPUTER READABLE

COMPUTER: IBM PC COMP

```

;
; OPERATING SYSTEM: PC-DOS/MS
;
; COMPILER: Borland Turbo Pascal 4.0

```

; CURRENT APPLICATION DATA:

;
FILING DATE:

CLASSIFICATION : PRIOR APPLICATION

APPLICATION NUMBER:

; ATTORNEY/AGENT INFORMATION:

; REGISTRATION NUMBER: 18,8

: TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

; TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 7:

SEQUENCE CHARACTERISTICS:

LENGTH: 3155 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: cDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2943

us-09-442-100-7

Query Match 19.1%; Score 760; DB 4; Length 3155;

Best Local Similarity 63.7%; Pred. No. 5.2e-188;

Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

Qy 1653 CACTCTGCTAATCTCAGCCTTCTGCTACACAGTCTACTGCAATACACAGCTCCTATT 1712
 Db 985 CAACCTGAGCCCTCACTGCGCGCCCAACACAGCTCACCGCGGTGAGCGCGCCACACATC 1044
 Qy 1713 CAACAGCCTGTGAAGATGCTGTATTAACACAGCTACAGACTGCTTTAGCACCT 1772
 Db 1045 CTTACCCCTGTGAAGAGCGTGTGCTGCGGCCCGAGCCCGAGACCGCGTGGGCCCC 1104
 Qy 1773 ACACACCCCTCTTGGATACACAGCAATTCAAACTGTTCAAACCCAGTCTCTTTCTCTGAG 1832
 Db 1105 TCGACCCCGCTGGTGGCTGCGCCACACAGCACTGCCACTGAGACCTGGAGACGAG 1164
 Qy 1833 GGAACCGCTTCAAAATGTGACTGTGATGCCACCTTGTGCTGAAGCTCCAAATATCAAGGA 1892
 Db 1165 GAGGACGCGCAGGCCCAACACCGCTGTGATGTGACTATGCGGCTCCGAGCGCAGGTGC 1224
 Qy 1893 CCACACACCCCTACCCAAACATCTGCTGCACCAAAACCCATCTGTCTCTCCATACGAG 1952
 Db 1225 CCACCGCTCTGCTATCAAGACACTTGTGCTGCCCACTAAGTCTGACGACTACAGCTG 1284
 Qy 1953 TCAATCAGTAAGCTTAGCAAGAGAGATCAGCAAGCTTGCCTCAAGAGAGATGAGAGTGA 2012
 Db 1285 GACCTGGACGCTGTGCACCACTGTGCAGCAGAGTCTGCGAGGGGGCACTGATAGAC 1344
 Qy 2013 AAGAGTTA-----TGAAATGTGTAGTGGGGATGAAGAGGAGGAGGAGGAGTCT 2057
 Db 1345 GGGAGTGAAGAGGACCAAGAGGTGCGAAGGAGGAGCAAAAGTGGCAGAGCAAAAGAGCAG 1404
 Qy 2058 ATTACAACCTTCACCTATTACTTGTAGGAAAAACAAGAGATGAAGAGGAGGAGGAGTCT 2117
 Db 1405 ATTACAGACTCCCGGCTGCTGCTGCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
 Qy 2118 CGTATTCAAGTTATTCCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAT 2177
 Db 1465 CGCATCAAGAGTTACTCCCTTATGCCCTTCAATTTCTTATGGAGCAACAGCTGGAGAA 1524
 Qy 2178 GTACTCAAAATCTCATCAGCAGCGTCTACATCGTAAAAACAATTAAGAAATGAATGATG 2237
 Db 1525 GTACTCAAAATCTCATCAGCAGCGTCTACATCGTAAAAACAATTAAGAAATGAATGATG 1584
 Qy 2238 CGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGGAAAAAGAA 2297
 Db 1585 AAAGCTGGCTCTGTGAGGCGCAGCAGCAGATGAGGAGATCTCTACCAAGAGGAG 1644
 Qy 2298 TCTAATATACATCCGCTTTAAAGGGCTAAATGGACAAGTCTATGTTTGTGAAGTAAG 2357
 Db 1645 TCTAATATACATCCGCTTTAAAGGGCTAAATGGACAAGTCTATGTTTGTGAAGTAAG 1704
 Qy 2358 ACACATAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAATAGATACCTAAGCT 2417
 Db 1705 ACTCTAGGCATCGGTCCTTTGGGAAAGTGTGCCCTGTGTAAGCTGGACACTCAGCT 1764
 Qy 2418 TTGTATGAACAAAACTCTTCGAAGAAAGATGTTCTTCTTCGAAATCAAGTCGCTCAT 2477
 Db 1765 CTGTAGCCATGAAGACTCTCAGGAAGAGAGTGTCTTGAACCGGAATCAAGTGGCCCAT 1824
 Qy 2478 GTTAAGGCTGAGAGAGATATCTGCTGAAGCTGAAGTGAATGAATGAGTGTCTATAT 2537

Db 1825 GTCAAGGCTGAGAGGACATCTCTGGCTGAAGCAGACAATGAGTGGTGTCAAACTCTAC 1884
 Qy 2538 TATTCAATCAAGATAAGGACAATTTATCTTTGTAATGACTACATTTCTCTGGGGTGAT 2597
 Db 1885 TACTCTTTCCAGGACAGGACGCCCTGTACTTTGTGATGACTACATACAGCGGGGAT 1944
 Qy 2598 ATGATGAGCCTATTAAATTAAGATGGCATCTTTCCAGAAAGTCTGGCAGATTTACATA 2657
 Db 1945 ATGATGAGCCTGCTGATCAGGATGAGGCTCTTCCCTGAGCACCTGGCCGCTTCTACAT 2004
 Qy 2658 GCAGAACTTACCTGTGCGAGTTGAAAGTGTCAATAAATGGGTTTTTATTCATAGATATT 2717
 Db 2005 GCAGAGTTGACCTGGCCATTGAAAGTGTCCACAAGATGGGCTTTATPCCACCGGACATC 2064
 Qy 2718 AAACCTGATAATATTGATGATCGTGTGATCAATTAATTAATGACTGACTTTGGGCTC 2777
 Db 2065 AAGCCTGACAACATACATCGACCTGGATGGTCAATTAAGCTGACAGATTTTGGCCTC 2124
 Qy 2778 TGCACCTGCTTCAGATGGACACACAGATTTTAAGTACTATCAGAGTGTGACCATCCAGG 2837
 Db 2125 TGCACCTGATTCAGTGGACTCAATTTCCAAGTACTACCAAGAAAGGAAACCAATGAGA 2184
 Qy 2838 CAAGATAGATGATTTCAATTAAGTAAATGGGGATCCCTCAAGCTGTGCTGAGAGAC 2897
 Db 2185 CAGGACAGCATGGAGCCCGGTGACCTCTGGACAGATTTTCCAACCTGTGCTGTGGAGAC 2244
 Qy 2898 AGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCAGCACAGGAGTGTCTAGACATCTCT 2957
 Db 2245 AGGTTAAAGACCTTGGAGCAGAGAGGCGCCAGAACAGCAGAGGTCCTTGGCACATCT 2304
 Qy 2958 TTGTTGGGACTCCCAATATTATTCACCTCAAGTGTGTGAGCAAGAGATACACAGAC 3017
 Db 2305 CTGTGCGGACACCAATTTACATCGCTCCGAGGTCTTCCGCAAAAGGAGTACACAGAC 2364
 Qy 3018 TTGTGCTATTGGTGGAGTGTGTTTATTTCTTTTGAATGTTGTGGTGGACACCTCTCT 3077
 Db 2365 CTCTGTGACTGGTGGAGCGTGGTGTGATTTCTCTTTGAGATGCTGTGGTGGGACGCCCT 2424
 Qy 3078 TTCTTGGCAACCAACCATTTAGAAACACAAATGAAGTGTATCAACTGGCAACATCTCTT 3137
 Db 2425 TTCTTGGCCCCCCCCACACAGACGACGCTGAAGTGTATCACTGGAGAGCACGCTG 2484
 Qy 3138 CACATTCACACAAAGCTAACTCAGTCTCTGAGAGTCTCTGATCTTATTTAACTTTGC 3197
 Db 2485 CATATCTCTACGCAAGTGAAGCTCAGGCTCAGGCTGAGGCCGAGACCTCATCAGAACTGTC 2544
 Qy 3198 CGAGGACCGAAGATCGCTTAGGCAAGAAATGCTGTGATGAAATAAAAGCTCATCCATTT 3257
 Db 2545 TCGCGGCTGACTCGCGCTGGGAGGATGGGAGATGAGCTCAAGGCACACCCGCTTC 2604
 Qy 3258 TTTAAACAAATTTGACTTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATCATCTCTAAA 3317
 Db 2605 TTCAACACCATCGACTTTTCCCGTGACATCCGAAAGAGGCTGCACCTCAGTCCGCCACC 2664
 Qy 3318 ATCACACACCAACAGATACATCAAAATTTTCTGATCTCTGATGATAAATTTATGAGT 3377
 Db 2665 ATCAGCCACCCCTGAGACCTCCAAATTTTACCCGTTGATGAAGAAAGCCCTCGGCAC 2724
 Qy 3378 GATGATAACGAGGAGAAATGTAATGACACTCTCAATGATGATGATAAATAATGGAAG 3437
 Db 2725 GAGGCCAGC---GGAGAGAGCGCAAGGCTTGGGACAGGCTGGGCTCCCGCAGCAGCAAG 2781
 Qy 3438 CATCTGAACATGCTATCTATGAATTTACCTTCCGAAGGTTTTTTTGTGATGACAAATGGCTAC 3497
 Db 2782 CATCCAGAGCAGCGCTTCTATGAGTTCACCTTCGAGGTTCTTCGATGACAAAGGCTAT 2841
 Qy 3498 CCATATTAATTTCCGAAGCC 3517
 Db 2842 CCTTCCGTTGCCGGAAGCC 2861

RESULT 4

US-09-509-902A-15

; Sequence 15, Application US/09509902A
; Patent No. 6387676

; GENERAL INFORMATION:

; APPLICANT: Virca, Duke

; APPLICANT: Bird, Timothy A.

; APPLICANT: Anderson, Dirk M.

; APPLICANT: Marken, John S.

; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions

; FILE REFERENCE: 2877-US

; CURRENT APPLICATION NUMBER: US/09/509,902A

; CURRENT FILING DATE: 1999-08-03

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 15

; LENGTH: 1961

; TYPE: DNA

; ORGANISM: Homo sapiens

; US-09-509-902A-15

Query Match 18.6%; Score 742.8; DB 4; Length 1961;
Best Local Similarity 66.2%; Pred. No. 1.2e-183;
Matches 1122; Conservative 0; Mismatches 557; Indels 15; Gaps 3;

QY 1872 GAAGCTCCAAACTATCAAGGACCAACACCCCTACCCAAAACATCTGCT-----GCACC 1926
DB 21 GGAGGCCACAGCGGAGGTGCCCGCTCCGCTACCCGAAGCACCTGCTGCTGCGCAGC 80
QY 1927 AAAACCCATCTGTTCTCCATAGGAGTCAATCAGTAGGCTAGCAAGAGGATCAG---- 1982
DB 81 AAGTGGAGCAGTACGACCTGGACAGCCCTGTGCGCAGGATGGACAGAGCCTCCGTGCG 140
QY 1983 ---CAAAGCTTCCCAAGAAAGATGAGAGTGAAGAGAGTTATGAAATGTTGATAGTGG 2039
DB 141 GSCCCCAAGACGCGCGGCGGCGACAGAGCGCGGAAAGCGCCAGGGGACAAAGCG 200
QY 2040 GATAAGAAAAGAAACAGATTACAACTTCCACTTATTACTTGTAGGAAAACAAAGAAAGAT 2099
DB 201 GGAAGGATAAAAGCAGATTACAGACTCTCCGTTCCCGTCCGCAAAAACAGCAGAC 260
QY 2100 GAAGCGGAAGGAATCTCGTATTCAAGTTATTCTCTCAAGCATTTAAATCTTTATG 2159
DB 261 GAAGAGACAGAGAGTACGATCAAGAGTACTCGCCATAGCCCTTTAAGTTCTTCATG 320
QY 2160 GAGCAACATGTAGAAATCTACTCAATCTCATCAGCAGCGTCTACATCGTAAAAACAA 2219
DB 321 GAGCAGCAGCTGAGAAATGTATCAAAACCTACCAGCAAGATTAACCGAGGCTGCAG 380
QY 2220 TTAGAGATGAATGATCGCGGTTGGATTTATCTCAAGATGCCAGGATCAAAATGAGAAAG 2279
DB 381 CTGGAGCAAGAAATGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGACAGATCGCGAAG 440
QY 2280 ATGCTTTGCCAAAAGAAATCTAATTACATCCGCTTTAAAGGGCTTAAATGGACAAGTCT 2339
DB 441 ATCTCTTACCAAGAAAGAGTCTAATTACACAGGTTAAAGAGGGGCCAAGATGAGCAAGTCT 500
QY 2340 ATGTTTGTCAAGTAAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTGTAGCAAGA 2399
DB 501 ATGTTTGTCAAGTCAAAACCTCGGGATCGGGTCTGGTGGAGAGTGTCCCTTGTCTGT 560
QY 2400 AAAGTAGATPACTAAGCTTTGTATGCAACAAAACCTTTTCGAAAAGAGATGTTCTTCTT 2459
DB 561 AAGGTGGACACTCAGCCCTGTACGCCATGAAGACCTTAAGGAAAAAGATGTCTCTGAAC 620
QY 2460 CGAATCAAGTCGTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACAATGAA 2519
DB 621 CGGAATCAGGTGGCCACCTCAAGGCGGAGAGGAGATCCTGGCCGAGGACACATGAG 680
QY 2520 TGGGTAGTTCGCTATATTATTCATTCGAAGATAGGACAATTTATCTTTGTAATGGAC 2579
DB 681 TGGGTGGTCAAACTCTTACTTCTTCCCAAGACAAGACAGGCTGTACTTTGTGATGGAC 740
QY 2580 TACATTCTCTGGGGGTGATATGATGAGCCCTATTAAATTAAGAATGGCATCTTTCCAGAAAGT 2639

RESULT 5.

US-09-509-902A-6

; Sequence 6, Application US/09509902A

; Patent No. 6387676

; GENERAL INFORMATION:

```

; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-509-902A-6

```

	Query Match	17.0%	Score 675.8	DB 4	Length 1498
	Best Local Similarity	67.2%	Pred. No. 3e-166		
	Matches 991	Conservative	0	Mismatches 472	Indels 12
	Gaps				
QY	1872	GAAGCTCCAAACTATCAAGGACACACACACCTACCCCAAAACATCTGCT-----GCACC	1926		
DB	23	GGAGGCCACAGCCGGAGTGCCCGCTTCGGCCCTACCCGACAGCCTGCTCTCGCGAGC	82		
QY	1927	AAAACCCATCTGTTCCTCCATACGAGTCAATAGTAAGCCTAGCAAGAGGATCAG-----	1982		
DB	83	AAGTCGGAGCAGTACGACCTGGACAGCCTGTGCGAGGCATGGACAGACGCTCCGTGCG	142		
QY	1983	---CCAAGCTTCCCAAGCAAGATGAGTGAAGAGACTTATGAAATGTGTGATAGTGGG	2039		
DB	143	GGCCCCACAGAGCCGAGCGGCGGCGCAGAGCGCCGCAAAAGCCGAGGGGACAAAGGC	2023		
QY	2040	GATAAAGAAAAGAACAGATTACAACCTTACCCTATTACTGTAGGAAAAACAAAGAAAGAT	2099		
DB	203	GGAAGGATAAAGACAGATTACAGACCTCTCCCGTTCGCGCCGCAAAAACAGCAGAGAC	262		
QY	2100	GAGAGGGAAGGGAATCTCGTATTCAAAGTATTCTCTCTCAAGCATTTAAATCTTTTATG	2159		
DB	263	GAAGAGAAGAGAGTACGCGCATCAAGAGCTACTCGCCATACGCTTTAAGTCTCTTCATG	322		
QY	2160	GAGCAACATGTAGAAAATCTACTCAAATCTCATCAGCAGCGTCTACATCGTAAAAACAA	2219		
DB	323	GACGACGCTGGAGAAATGTATCAAAACCTACCAGCAGAGAGTTAACGGAGCGCTCGAG	382		
QY	2220	TTAGAGAATGAATATGCGGGTGTGATTTATCTCAAGATGCCAGAGATCAAAATGAGAAG	2279		
DB	383	CTGGAGCAAGAAATGCCAAAAGCTGCTGTGAAGCTGACGAGGACAGATCGGAAG	442		
QY	2280	ATGCTTTGCCAAAAGAAATCTAATTACATCCGCTCTTAAAGGGCTTAAATGSGCAAGTCT	2339		
DB	443	ATCCTCTACCAGAAAGAGTCTAATTACACAGGTTAAAGGGCCAGATGACAGTCT	502		
QY	2340	ATGTTTGTGAAGATAAAGACACTAGGAATPAGGAGCATTTGGTGAAGTCTGTCTAGCAAGA	2399		
DB	503	ATGTTTGTCAAGATCAAAACCCCTGGGGATCGGTGCTTTGGAGAAAGTGTGCTTTGTGT	562		
QY	2400	AAAGTAGATACTAAGCGTTTGTATGCAACAAAACCTCTCGAAGAAAGATGTTCTTCTT	2459		
DB	563	AAGGTGACACTCACGCCCTGTACGCCATTGAGAGCCCTAAGGAAAAAGGATGTCTGAAC	622		
QY	2460	CGAAATCAAGCTCGCTCATGTTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACAATCAA	2519		
DB	623	CGGANTCAGTGGCCACAGTCAAGGCCGAGAGGACATCTCTGGCGAGGACAGCAATGAC	682		
QY	2520	TGGTAGTTCGTCTATATTATTCATPCCAAGATAAGGACAAATTTATCTTTGTATATGAC	2579		
DB	683	TGGTGTGTCAAACCTACTACTCTCTCCAAGACAAAGACAGCCTGTACTTTGTGATGGAC	742		
QY	2580	TACATTCCTGGGGGTGATATGATGACCTTAAATTAGNATGGGCATCTTTCCAGAAAGT	2639		
DB	743	TACATCCCTGGTGGGACATGATGAGCCCTGCTGTATCCCGATGGAGGTCTTCCCTTGAGC	802		
QY	2640	CTGGCAGATTTCTACATAGCAGAACTTACCTGTGCAAGTTGAAAGTGTTCATATAAATGGT	2699		

RESULT 6
US-09-328-111-26/c
Sequence 26, Application US/09328111
Patent No. 6262333
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/328,111
CURRENT FILING DATE: 1999-06-08
EARLIER APPLICATION NUMBER: US 60/088,801
EARLIER FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850

SOFTWARE: FastSeq for Windows Version 3.0

SEQ ID NO 26

LENGTH: 638

TYPE: DNA

ORGANISM: Homo sapiens

US-09-328-111-26

Query Match 15.3%; Score 609.8; DB 4; Length 638;

Best Local Similarity 98.4%; Pred. No. 3.1e-149;

Matches 627; Conservative 0; Mismatches 7; Indels 3; Gaps 1;

QY 2178 GTACTCAATCTCATCAGCAGCGTCTACATCGTAAACAAATAGAGAAATGAATGATG 2237
 DB 637 GTACTCAATCTCATCAGCAGCGTCTACATCGTAAACAAATAGAGAAATGAATGATG 578
 QY 2238 CGGTTGGATTATCTCAAGATGCCAGGATCAATCAGAAAGATGTTGCCCAAAAGAA 2297
 DB 577 CGGTTGGATTATCTCAAGATGCCAGGATCAATCAGAAAGATGTTGCCCAAAAGAA 518
 QY 2298 TCTAATACATCGTCTTAAAGGGCTTAAAGTGAAGTCTATGTTGTGAAGATAAAG 2357
 DB 517 TCTAATACATCGTCTTAAAGGGCTTAAAGTGAAGTCTATGTTGTGAAGATAAAG 458
 QY 2358 ACATAGGATAGGACATTTGGTGAAGTCTGTCTAGCAAGAAATAGATAGAGCT 2417
 DB 457 ACATAGGATAGGACATTTGGTGAAGTCTGTCTAGCAAGAAATAGATAGAGCT 398
 QY 2418 TTGTATGCAACAAACTCTTCAAGAAAGATGTTCTTCAAGAAATCAAGTCGCTCAT 2477
 DB 397 TTGTATGCAACAAACTCTTCAAGAAAGATGTTCTTCAAGAAATCAAGTCGCTCAT 338
 QY 2478 GTTAAGGCTGAGAGATATCTCGCTGAAGCTGACAATGAATGGTGTCTGTCTATAT 2537
 DB 337 GTTAAGGCTGAGAGATATCTCGCTGAAGCTGACAATGAATGGTGTCTGTCTATAT 278
 QY 2538 TATTCATCCAGATAAGGACATTTATATCTTTGTAATGACATACATTCCTGGGGTGAT 2597
 DB 277 TATTCATCCAGATAAGGACATTTATATCTTTGTAATGACATACATTCCTGGGGTGAT 218
 QY 2598 ATGATGAGCTATTAATAGATGGCATCTTCCAGAAAGCTGCGCAGATCTCATATA 2657
 DB 217 ATGATGAGCTATTAATAGATGGCATCTTCCAGAAAGCTGCGCAGATCTCATATA 158
 QY 2658 GCAGAACTTACCTGTGAGCTGAAAGTGTTCATATAAATGGTGTTCATAGAGAT 2714
 DB 157 GCAGAACTTACCTGTGAGCTGAAAGTGTTCATATAAATGGTGTTCATAGAGAT 98
 QY 2715 ATTAACCTGATAATTTTGTATGATCGTGATGGTCAATTAATGACTGACTTTGGC 2774
 DB 97 ATTAACCTGATAATTTTGTATGATCGTGATGGTCAATTAATGACTGACTTTGGC 38
 QY 2775 CTCTGACCTGGCTTCAGATGGACACAGATTTCTAAGT 2811
 DB 37 CTCTGACCTGGCTTCAGATGGACACAGATTTCTAAGT 1

RESULT 7

US-09-442-100-1

Sequence 1, Application US/09442100

Patent No. 6359193

GENERAL INFORMATION:

APPLICANT: Xu, Tian

APPLICANT: Tao, Wufan

APPLICANT: Wang, Welyi

APPLICANT: Zhang, Sheng

APPLICANT: Yu, Wan

TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS

TITLE OF INVENTION: GENES AND METHODS BASED THEREON

NUMBER OF SEQUENCES: 16

CORRESPONDENCE ADDRESS:

ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas

CITY: New York

STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/442,100
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/411,111
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 6523-003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5720 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1103..4402
 US-09-442-100-1

Query Match 11.8%; Score 469; DB 4; Length 5720;

Best Local Similarity 63.0%; Pred. No. 4.1e-112;

Matches 819; Conservative 0; Mismatches 455; Indels 27; Gaps 5;

QY 2088 ACAACAAGATGCAAGAGGGAATCTCGTATTCAAAAGTTATCTCTCAAGCATTT 2147
 DB 2984 AAGGAGAGGAGGAGGAGGCAAGGAGTTCCCGCATCAGGCAGTACTCGCGCAAGCTTC 3043
 QY 2148 AAATCTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGCTCTACAT 2207
 DB 3044 AAGTCTTCTATGGAGCAGCAGACATAGAGACGTGATCAAGTCTGATCGCCAGCAGCTAT 3103
 QY 2208 CGTAAAAACAATAGAGAAATGAATGATCGCGGTTGGATTATCTCAAGATGCCCCAGGAT 2267
 DB 3104 CGCAAGAATCAGCTGGAGAAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCCAATC 3163
 QY 2268 CAATGAGAAAGATGCTTTGCCCAAAAGAAATCTAATATACATCGCTTTAAAGGGCTAA 2327
 DB 3164 GAGATGAGAAATGCTGAACCAAAAGGAGACCAATACATTCGATTGAGCGCGCAAG 3223
 QY 2328 ATGGACAAGTCTATGTTGTGAAGATAAAGACACTAGGATAGGAGCTTTGGTGAAGTC 2387
 DB 3224 ATGGACAAGCATGTTCTGCAAACTGAGCCCATTTGGAGTGGTGCATTTGGCGAGGTA 3283
 QY 2388 TGTCTAGCAAGAAAGTAGATA---CTAAGGCTTTGTTATGCAACAAAACTCTTCGAAAG 2444
 DB 3284 ACCTGCTGAGCAAAATCGATACCTCGAACCATTTCTATGCGATGAAACCCCTGCGGAAA 3343
 QY 2445 AAAGATGTTCTTCTGCAATCAAGTCGCTCATGTTAAGCGTGAGAGAGATATCTTGGCT 2504
 DB 3344 GCGAGCTTCTCAAGCGGAATCAGTGGCACACGTCGAGGAGGAGGATATCTTCTCGG 3403
 QY 2505 GAAGCTGACAAATGAATGGGTAGTTCGCTCTATATTATTCATTCACAGATGAAGCAATTA 2564
 DB 3404 GAAGCCGACAAATGAATGGGTAGTTCGCTCTATATTATTCATTCACAGATGAAGCAATTA 3463
 QY 2565 TACTTTGTAATGAGCTACATTCCTGGGGTGATATGATGAGCCTATTATTAATGAGTGGC 2624

Fri Jan 17 11:16:54 2003

Db 3464 TACTTTGTGAGTACTATACACAGTGGTGTGATCTGATCTGCTGCTATCAAACTGGGC 3523
QY 2625 ATCTTTCCAGAAAGTCTGGCAGGATCTACATAGCAAACTTACCTGTGCGAGTTGAAGT 2684
Db 3524 ATTTTCGAGGAGAACTGGCCAGATTTCTACATCCGCGAGTCACTTCCGCGTGGACAGC 3583
QY 2685 GTTTCATAAAATGGGTTTATTCATAGAGATATTAACCTGATATATTTGATGATCGT 2744
Db 3584 GTTCACAAATGGGCTTCATTCACAGACATCAAGCTGACACATCACTCATCGATAGG 3643
QY 2745 GATGTCATATTAATTCAGTCTGCTTGGCTCTGCACTGGCTTCAGATGGACACAGAT 2804
Db 3644 GACGGACACATAAAGCTCACCGACTTTGGCCCTGTGACGCGGATTCGATGGACACAAC 3703
QY 2805 TCTAAGTACTATC---AGAGTGGTGACCATCCACGGCAAGATAGCATGATGATTTTCAAT 2861
Db 3704 TGAAGTACTACAGGAGACGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAG 3763
QY 2862 GAATGGGGGATCCCTCAAGCTGTGATGTGGAGACAGACTGAAGCCATTAAGACGGAGA 2921
Db 3764 GAATACTCCGAGAAC-----GGACCGAAGCCACCGTCTCGAGAGCGGA 3808
QY 2922 GCTGACCGCAGCAGCAGGATGCTTAGCACATTTCTTGGTGGGACTCCCAATATATT 2981
Db 3809 CGGATGGCGGATCACAAGAGTCTTGGCCACTCGTGGTGGGACCCCGAAGCTACATA 3868
QY 2982 GCACCTGAAGTGTGCTAGCAGGATACACAGATGCTGTGATTTGGTGGAGTGTGGT 3041
Db 3869 GCTCCGAGGTGTGGAGAGAGTGGGTACACGACACTGTGGACTACTGGAGCGTGGGC 3928
QY 3042 GTTATTTCTTTTGAATATGTTGGGACACACTCTTTCTTGGCACAACACACCATTAGAA 3101
Db 3929 GTCATCTTTTACGAGATGCTGGTGGTCAAGCGCCCTTTCTGGCCAAACAGTCCGCTGGA 3988
QY 3102 ACACAAATGAAGTATCACTGGCAACATCTTCAACATCCACCAACAGTAACATC 3161
Db 3989 ACGCAACAAAGGTCATCACTGGGAGAAACCGTGCATATTCGCGCAGGCGGAGTTA 4048
QY 3162 AGTCTGAAGCTCTGATCTTATTTATTAATTTAACTTGGCGAGGACCCGAGATCGCTTAGGC 3221
Db 4049 TCCCGGAGGTACGAGCTGATAAGAGGCTCTGTGCTGGCTGACAAAGCGCTTGGC 4108
QY 3222 AAGATGCTGTGATGAATAAAGCTCATCCATTTTAAACAAATGACTTCTCCAGT 3281
Db 4109 AAGA---CGCTGGAGGAGGTCAAGACCGACACTTCTTCAAGGCGCTCGACTTGC---G 4162
QY 3282 GACCTGAGACAGCTCTGCTTCAATTCCTTAATTCACACACACCAACAGATACATCA 3341
Db 4163 GACATGCGGAACGCAAGAGCGCCCTTACATACCGGAAATCAAGCAACCGGACATCC 4222
QY 3342 AATTTGATCTGTTGATCTGATATTAATTTATGAGTATGA 3382
Db 4223 AACTTTGATCCGCTGATCCGGAAGAGCTGCGCTGATGA 4263

RESULT 8
US-09-328-111-66
; Sequence 66, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)...(678)
; OTHER INFORMATION: n = A,T,C or G
US-09-328-111-66
Query Match 11.4%; Score 452.8; DB 4; Length 678;
Best Local Similarity 90.1%; Pred. No. 2.4e-108;
Matches 548; Conservative 0; Mismatches 50; Indels 10; Gaps 6;
QY 2180 ACTCAATCTCATCAGCAGCGTCTACATCGTAAAAACAATTAGAGATGAATGATGCG 2239
Db 1 ACTCAATCTCATCAGCAGCGTCTACATCGTAAAAACAATTAGAGATGAATGATGCG 60
QY 2240 GGTGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTCCAAAAAGATC 2299
Db 61 GGTGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTCCAAAAAGATC 120
QY 2300 TAATACATCCGCTTAAAGGGCTAAATGGACAGTCTATGTTTGTGAAGATAAGAC 2359
Db 121 TAATACATCCGCTTAAAGGGCTAAATGGACAGTCTATGTTTGTGAAGATAAGAC 180
QY 2360 ACTAGGAATAGGAGCTTGGTGAAGTCTGTCTAGCAAGAAAGTACTAAGGCTTT 2419
Db 181 ACTAGGAATAGGAGCTTGGTGAAGTCTGTCTAGCAAGAAAGTACTAAGGCTTT 240
QY 2420 GTATGCAACAAAACCTCTTCAAGAAAGATGTTCTTTCGAAATCAAGTCGCTCATGT 2479
Db 241 GTATGCAACAAAACCTCTTCAAGAAAGATGTTCTTTCGAAATCAAGTCGCTCATGT 300
QY 2480 TAAGCTGAGAGAGATATCTTGGTGAAGTCTGAGTGAAGTGAATGGTCTATATTA 2539
Db 301 TAAGCTGAGAGAGATATCTTGGTGAAGTCTGAGTGAAGTGAATGGTCTATATTA 360
QY 2540 TTCATTCCAAGATAGGACAATTTATATCTTTGTAATGGACTACATTCCTGGGGTGATAT 2599
Db 361 TTCATTCCAAGATAGGACCATTATTCCTT--GTAATGGCTACATTCCTNGGGTGATAT 418
QY 2600 GATGAGCCTTATTAATAGAAATGGGATC-TTTCAGAAAGTCTGCACGATTCATCAT-- 2656
Db 419 GAAGAGCCCATTAATANAATGGGCTCTTTTCCAGAAAGGCTNGCACCATCTACCTTA 478
QY 2657 AGCAGAACTTACTGTG-CAGTTGAAAGTGTTCATAAAA--TGGGTTTTTATTCATAGAGA 2713
Db 479 GCAGAACTTACTGTGCGCCNGCTTGAAGTGGTCTTAAATGGGCTTAAATTCCTTAGAGA 538
QY 2714 TATTAACCTGATAATATTTTTCATGATC--GTGATGCTCATATTAATTAAGTACTG 2771
Db 539 TTTTAACTGGGATAATATTTGANTGGACCGAAGGCCCTTATTAATTAAGTACTGCTTT 598
QY 2772 GGCCTGTG 2779
Db 599 GGCCTTNG 606

RESULT 9
US-08-788-989-11
; Sequence 11, Application US/08878989
; Patent No. 5885803
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga

Db	422	GATATGCTTGA	AAAAAGAGCAGCGTGGCCCATATCCGAGCAGAAAGAGATATTTTGGTAGAA	481
QY	2508	GCTGCAAAATGA	ATGGGTAGTTCGCTCATATATTATTTCATTTCCAAAGATAAAGACAATTTATAC	2567
Db	482	GCAGATGGTGC	TGGTGGTGAAGATGTTTTACAGTTTTTCAGGAYAAAGAGAAATCTTTAT	541
QY	2568	TTTGTAAATG	ACTACATCTCTGGGGGTGATATGATGACCTTATTAAATTAGAATGGGCATC	2627
Db	542	CTAATCATG	AAATTTCTCCCTGGAGTGCATGATGACATTTGCTTAATGAAGAAGACACC	601
QY	2628	TTTCCAGAAG	CTCTGCGAGGATTTACATACGACAGACTTACCTGTGCAGTTGAAAGTGT	2687
Db	602	TTGACAGA	AGAGAAACACAGTCTTACATTTTCAGAGACTGTTCTGGCAATAGATGGCAT	661
QY	2688	CATAAAAT	GGGTTTTATTTCATAGACATATTAACCTGATATAATTTTGAATGATCGTAT	2747
Db	662	CACCACTT	GGGTTTCATCCATCGGGATATTAGCCAGACAACCTTTTATTGGATGCCAAG	721
QY	2748	GGTCATAT	TAAATTTGACTGACTTTGGCCCTCTGCACCTGGCTTCAGATGGACACACGATTC	2807
Db	722	GGTCATG	TAAATTTATCTGATTTTGGTTCATGTACGGGATTAAGAAGCTCACAGACT	781
QY	2808	AGTACTAT	CAGAGTGGTGACCATCCAGCGCAAGATAGCATGGATTTCACTAATGAATGG	2867
Db	782	GAATTTAT	-----AGAAATCTCACACAACCCACCAAGTG	818
QY	2868	GGGATCCCT	CAAGCTGTCGATGTGGAGACAGACTGAAGCCATTTAGAGCGGAGAGCTGCA	2927
Db	818	ACTTCTCAT	TTTCAGAACATGAACCTCAAGAGAGAAAGCAGAACTTGGAAAGAAACAGGA	878
QY	2928	CGCCAGC	ACCACGCGATGCTAGCACATCTTTTGGTTGGGACTCCCAATTTATTTGCACCT	2987
Db	878	GACAA-----	CTGGCATATTCACAGTTGGGACACACGATTTACATTTGCTCCA	925
QY	2988	GAAGTGTG	CTACGAACAGGATACACAGTTGTGTGATTTGGTGGAGTGTGGTGTATT	3047
Db	926	GAAGTATTC	ATCGACACTGGTTTACAACAAAATTTGTGACTGGTGGTCTTTGGGAGTGAT	985
QY	3048	CTTTTGAAT	GTTGGTGGGACCAACCTCCTTTCTTTGGCACAACACCAATTTAGAAACACAA	3107
Db	986	ATGATATGA	AAATGCTAATAGGATATCCACTTTCTGCTCTGAACACACCTCAAGAAACATAC	1045
QY	3108	ATGAAGGTT	ATCAACTGGCAACATCTTTCACATCTTCCACCAAGCTAAACTCAGTCTCT	3167
Db	1046	AGNAAAGT	ATGATGAAAGAACTGAAAGAAACTCTGATTTTCTCCAGAGTTACCTATATCTGAG	1105
QY	3168	GAAGCTTC	TGATCTTTATTAAACTTTTGGCAGGACCCGAGAGATGCGTTAGGCAAGAAT	3227
Db	1106	AAAGCC	AAAGGACTTAATCTCAGATTTTGTATTGATCTGAAACACAGATTTGGAATAGT	1165
QY	3228	GGTGTGAT	GAAATAAAGCTCATCCATTTTTTAAACAAATTTGACTTCTCCAGTGACCTG	3287
Db	1166	GGAGTAGA	AGNAATAAAGGTCATCCCTTTTTTGA---AGTGTGACTGGGAGCACATA	1222
QY	3288	AGACAG	CACTGCTTCTATACATTCCTAAAACTCACACACCACACAGATACATCAAAATTTT	3347
Db	1223	AGGGAAG	CCAGCAGCAATCCCTATAGAAATCAAAAGCATTTGATGATACCTTCAATTTT	1282
QY	3348	GAT	3350	
Db	1283	GAT	1285	

RESULT 10
US-09-272-796-11
; Sequence 11, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti

```

QY 3168 GAAGCTCTCTGATCTCTTATTATTAACCTTTGCGGAGGACCCGGAAGATCGCTTTAGGCAAGAAT 3222
      |||||      |||      |||||      |||||      |||||      |||||      |||||
Db 1106 AAAGCAAGGACTTAATTCTCAGATTTTGTATTGATTTGAAAAACAGAAATTGGAATAGT 1165

QY 3228 GGTGCTGATGAATAAAAGCTCAATCCATTTTTTAAACAAATTGACTTCTCCAGTGACCTG 3287
      |||||      |||||      |||||      |||||      |||||      |||||      |||||
Db 1166 GGAGTAGAAGAATAAAAGGTCATCCCTTTTTTGA--AGGTGTGCGATGGGAGCACAATA 1222
      |||||      |||||      |||||      |||||      |||||      |||||      |||||

QY 3288 AGACAGCAGCTGCTTCATACATCTCTAAAATCACACACCCACAGAGATACATCAAAATTTT 3347
      |||||      |||||      |||||      |||||      |||||      |||||      |||||
Db 1223 AGGGAAGGCCAGCAGCAATCCCTATAGAAATCAAAAGCATTTGATGATACTTTCAAATTTT 1282
      |||||      |||||      |||||      |||||      |||||      |||||      |||||

QY 3348 GAT 3350
      |||||
Db 1283 GAT 1285

RESULT 10
US-09-272-796-11
; Sequence 11, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti

```


Db 845 CTTCGATGGAGATTTTGGATCTTAAAGTAAATAGCAGAGGAGCATTTGGTGAGGTA 904
Qy 2388 TGTCTAGCAAGAAAGTAGATACTAAGGCTTTGTATGCACACAAAACCTCTTCGAAAGAAA 2447
Db 905 CGGCTTGTTCAGAAAGAAATACGGGACATGTGTATGCAATGAAATACCTCCGTAAGCA 964
Qy 2448 GATGTTCTTCTTCAAAATCAAGTCGCTCATGTTAAAGCTGAGAGAGATATCCTGGCTGAA 2507
Db 965 GATATGCTTGAAGAAAGAGAGGTTGGCCACATTCGTGGGAGCGTGACATTTAGTGGAG 1024
Qy 2508 GCTCAGCAATGAATGGGTAGTTCGTCTATATTTATTCATTCCTCAAGATAAGGACATTTTATAC 2567
Db 1025 GCAGACAGTTTCTGGGTGTGCAAAATGTTCTATAGTTTTCAGGATAAGCTAAACCTCTAC 1084
Qy 2568 TTTGTAATGGACTACATTCCTGGGGGTGATGATGAGCCTTAATTAATAGATGGGATC 2627
Db 1085 CTAATCATGGAGTTCTGCTGGAGGGGACATGATGACCTTTGTTGATGAAAAGACACT 1144
Qy 2628 TTTCAGAAAGCTGGGACGATTTACATAGCAGAACTTACCTGTCAGTTGAAAGTGT 2687
Db 1145 CTGACAGAAAGAGAGACTCAGTTTTATATAGCAGAAACAGTATTAGCATAGACTCTATT 1204
Qy 2688 CATAAATGGGTTTTATTCATAGAGATATTAACCTTGATATATTTTATGATCGTGAT 2747
Db 1205 CACCAACTTGGATTCACACAGAGACATCAACACGACAACTTCTTTTGGACAGCAAG 1264
Qy 2748 GGTCAATTAATTAAGTACGACTTTGGGCTCTGCACTGGCTTCAGATGCACACAGATTC 2807
Db 1265 GGCCATGTGAACCTTTCTGACTTTTGGCTTTTGCACAGACTGNAAAAAGCACATAGGACA 1324
Qy 2808 AAGTACTATCAGAGTGGTGACCATCCAGGCAAGATAGCATGATTTTCAGTAATGAATGG 2867
Db 1325 GAATTTTATAGAACTCAACACAGCTCC-----CCAGTGATTCACITTTCCAGAAC 1378
Qy 2868 GGGATCCCTCAAGCTGTGATGTGGACAGACTGAAGCCATTAGCGGAGAGCTGCA 2927
Db 1379 ATGAATTCAAAAG-----GAAAGCAGAAAACCTGGA 1409
Qy 2928 GCCAGCAGCAGCGATGTCTAGCAGATTTCTTGGTGGGACTCCCAATATATATGACCT 2987
Db 1410 AAAGA-AATAGACGTCAGCTAGCTTCCACAGTAGGACACTCTGACTATTTGCTCT 1458
Qy 2988 GAAGTGTGCTACGAACAGGATACACAGATTTGTGATTTGGTGAGTGTGGTGTATT 3047
Db 1469 GAGGTGTTTCACGACACCGGTTACAAAGCTCTGTGATTTGGTGTGCTTTGGGCTGATC 1528
Qy 3048 CTTTGTGAATGTGGTGGGACACCTCTTCTTGGCACAACACCATTAGAACACAA 3107
Db 1529 ATGTATGATGCTCATCGGCTACCCACTTCTGTTCTGAGACCCCTCAAGAGACATAT 1588
Qy 3108 ATGAAGGTATCAACTGGCAACATCTCTTCACATTTCCACCAAGCTTAAACTCAGTCT 3167
Db 1589 AAGAGGTGATGACTGGAAGAAACTTTGACTTTTCTCCAGAGTTCCCATCTCTGAG 1648
Qy 3168 GAAGCTTCTGATTTATTTATTAATTTGGCGAGAGCCCGAGATCGCTTAGGCAAGAT 3227
Db 1649 AAGCCAAAGATCTAATTTGAGTCTGCTGTAATGGGAACATAGAAATTTGGAGCTCT 1708
Qy 3228 GTGCTGATGAATTAAGCTCATCCATTTTAAACAATTTGACTTCTCCAGTCAAGCTG 3287
Db 1709 GAGTGTGAGGAATAAAGTAATCTTTTGTGAGGCGTTGA---CTGGGAACATATC 1765
Qy 3288 AGACAGCAGCTTCTGCTTACATATTTCTTAAATACACACCCCAACATACATCAAAATTTT 3347
Db 1766 AGAGAGAGACCTGCTGCAATATCTATTGAAATCAAAAGCATTGATGATCACTCAAACTTC 1825
Qy 3348 GAT 3350
Db 1826 GAT 1828

RESULT 13

US-09-588-256-1
; Sequence 1, Application US/09588256
; Patent No. 6291665
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas
; APPLICANT: Flavler, Albert
; APPLICANT: Gates, Krista
; APPLICANT: Wendland, Juergen
; APPLICANT: Ayad-Durieux, Yasmina
; APPLICANT: Dietrich, Fred
; APPLICANT: Philippsen, Peter
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-30908A
; CURRENT APPLICATION NUMBER: US/09/588,256
; CURRENT FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Ashbya gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)..(2160)
US-09-588-256-1

Query Match 4.3%; Score 173.2; DB 4; Length 2160;
Best Local Similarity 56.2%; Pred. No. 2.7e-35;
Matches 325; Conservative 0; Mismatches 253; Indels 0; Gaps 0;
Qy 2243 TGGATTATCTCAAGATGCCAGGATCAANTGAGAAACATGCTTTGCCAAAAGAAATCTAA 2302
Db 828 TGGTGTCTCTGAAGAGAGAAGACACAGACGCTTTCTTCTACTGGGTAAAGAGGAGTCGCA 887
Qy 2303 TTACATCCGCTTAAAGGGCTAAATGGCAAGCTATGTTGTGAAGATAAAGACACT 2362
Db 888 GTTCTCGCTTGGTAGGACACGCGCTATCCCTGGAGATTTCCACACTGTTAAAGTCAT 947
Qy 2363 AGGAATAGGAGCAATTTGGTGAAGTCTCTAGCAAGAAAAGTAGATAGTAAAGCTTTGTA 2422
Db 948 AGAAAGGGTGCATTCGTGTAGGTCCTGCTCGTGCAGAAAGAAAGATACCCGGTAAATATA 1007
Qy 2423 TGCAACAAAACCTTTCGAAGAAGAGATGTTCTTCTCGAAATCAAGTCGCTCATGTAA 2482
Db 1008 CGCTATGAAGACATGTTTAAATCAGAAATGTACAAGAGATCAANTAGCCACGTCAA 1067
Qy 2483 GGCTGAGAGAGATATCCTGCTGAAGCTGACAAATGAATGGGTAGTTCGTCTATATTTATTC 2542
Db 1068 GGCCGAGAGGATGTGTGGCCGGAAGCGACTCTCCGTGGGTGCTGCTGTATATCTATTC 1127
Qy 2543 ATTCAGATGAAGACAAATTTATCTTTGTAATGGACTACATTCCTGGGTGATATGAT 2602
Db 1128 TTTCAGATGCCAGTACCTATATCTTGATCATGGAATTTTGGCCGGTGGTACCTGAT 1187
Qy 2603 GAGCTATTAATAGATGGGCTCTTTCCAGAAAGTCTGCGACGATTTCTACATAGCAGA 2662
Db 1188 GACCATGTTATCAGGTGGCAGATATTCACCGAGGACGTCACAGATTTCTACATGGCGGA 1247
Qy 2663 ACTTACCTGTCAGTGAAGTGTTCATAAATGGGTTTTTATTCATAGAGATNTAAACC 2722
Db 1248 GTGTATCTCGCAATTTGAGGCTATACACAAGCTGGGCTTTTATTCATAGAGATATCAAGCC 1307
Qy 2723 TGATAATATTTTGTGATGCTGATGCTGATTAATTAATTTGACTGACTTTGGCTCTGCAC 2782
Db 1308 GGATAACATTTCTGATCGACATCAGGGGTACATCAACTTTCCGACTTCGGTCTGCTCAC 1367
Qy 2783 TGGCTTTCAGATGGACACACGATTTCTTAAGTACTATCAGA 2820
Db 1368 AGGTTCCACAAAAGGCATGACTCCAACTACTACAAGA 1405

RESULT 14
US-08-860-150-1


```
; Sequence 1, Application US/08860150B
; Patent No. 5981205
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: PCT/EP95/05052
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: 94810746.1
; EARLIER FILING DATE: 1994-12-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2101
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (132)..(1499)
; US-08-860-150-1

Query Match      4.3%; Score 172.4; DB 2; Length 2101;
Best Local Similarity 49.9%; Pred. No. 4.3e-35;
Matches 539; Conservative 0; Mismatches 511; Indels 30; Gaps 3;

QY 2197 AGCGTCTACATCTATAAACAATAGAAATGAATGATGCGGGTTGGATTATCTCAAG 2256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 AGCGAAGCAGCGCCCTCGCAAGCTCAGGCTCAGCTGAAGGACGAGAGCTTCGCGAGG 321
QY 2257 ATGCCAGCATCAATAGAGAGATGCTTTGCCAAAAGAACTAATACATCCGCTTAA 2316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 CGCAGCGCCAGGAGAGCGCTCTGCAGCATGCCCAAGAGGAGCGAGTATCTCCGGCTGA 381
QY 2317 AAAGGCTAAAATGGCAAGTCTATTTCTGTAAGATAAAGACACTAGGAATAGGAGCAT 2376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 ACGATTGCGCTCGGTGGAGGACTTTGAGGCCCTCAAGATCATCGGCGCGGT 441
QY 2377 TTGGTGAATCTGTCTAGCAAGAAATAGATAGTAAAGCTTTGTATGCAACAAAACCTC 2436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 TCGTGAAGTGGCTTTGGTGCAGAAAAGGACACTGGACATGTGTGCGCCCATGAAGGTGC 501
QY 2437 TTCGAAGAAAGATGTTCTTCTTCGAAATCAAGTCTCATGTTAAGGCTGAGAG---AG 2493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TCGGCAAGCGGCATCTCTGGAAGAGGAGGAGGTGGCACACGTACCGCGGAGGCTGTCG 561
QY 2494 ATATCTGCTGAAGCTGACAATGAATGGTAGTTCTGCTCTATATTATTCATTCCAAGATA 2553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 ATGCTCTGCTGAGGCGGATCATCAGTGGTGTGAAGATGTACTACAGTTTCCAGGATC 621
QY 2554 AGCAAAATTAATTTTGAATGGACTACATCTCTGGGGGTGATATGATGAGCCTATTAA 2613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 CGGTCAATTTATTTGATAATAGAGTCTTTCCTGGTGGTGATATGATGACGCTTTTAA 681
QY 2614 TTAGAATGGCATCTTTCCAGAAAGCTGCGAGATTTCTACATAGACAGAACTTACCTGTG 2673
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 TGAAGAGCAGCGCTATCCGAGGAGGCGACACAGATTTCTATATCATGACGCGCATGG 741
QY 2674 CAGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGATATTAACCTGATAATATT 2733
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 CGATCATTTTATTCACAACTCGGTTTCATACACAGGATATACAGCCGATACCTTGC 801
QY 2734 TGATTGATCGTGAATGATATTAATTAATGACTGATTTGGCCCTCTGCATCGGCTTCAGAT 2793
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 802 TGTGAGCGCGGAGGCGCATCTGAAGCTCTCCGACTTCGCACTGTGCACTGGCTTAAAGA 861
QY 2794 GGACACAGATCTTAAGTACTATCAGATGGTGACCATCCAGCGGCAAGATAGCATGGATT 2853
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 862 AGTCGATCGAAGACAGACTTTTATTCGGGACTTTGTGCGAGGC-----GAAACCATCCGAT 914
QY 2854 TCAGTAATGAATGGGGGGATCCCTCAAGCTGTGCGATGTGAGACAGACTGAAGCCCATTAG 2913
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```
Db 915 TTTATAGGCACGTGCGCCAGTCCGATGGACTCCAAAGCGAGTCCCGAGTGTGGAAGCGA 974
QY 2914 AGCGGAGAGCTGCACGCCAGCAGCAGGATGCTTAGCACATTTCTTTGGTTGGACTCCCA 2973
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 975 AATCGAGCGCGCTCGCCTACAGCACC-----GTGGGAACGCCGG 1014
QY 2974 ATTATATTGCACCTGAAGTGTCTTACGAAACAGGATACACACAGTTGTGTGATTGGTGA 3033
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1015 ACTATATTGCACCGGAAGTATTTCTGCAGACTGGCTACGAGCCGCTGGACTGGTGGT 1074
QY 3034 GTGTTGGTGTATTCTTTTGAATGTGTGGGACAACCTCTCTTTGGCACAACAC 3093
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1075 CCCTGGGAGTCATCATGTACGAAATGCTGATGGCTATCTCCATCTGCTCGGACAATC 1134
QY 3094 CATTAGAAACACAATGAAGTTATCAACTGGCAAAACATCTTTCATCTCCACCACAAG 3153
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1135 CCAGGACACACTACCGCAAGGTGATGAACCTGGCGGAGAGCGGTATTTCCCCCAAGAG 1194
QY 3154 CTAAACTCAGTCTCTGAAGCTTCTGATCTTATTATAAATTTTCCGAGGACCCGAAGATC 3213
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1195 ATCCCATATCGGAGGAGGCCAAGGACGATCATCACTTCTGCTGGAGGCCGATCGCC 1254
QY 3214 GCTTAGCGAAGATGCTGCTGATGAATAAAGCTCATCCATTTTAAACAATTTGACT 3273
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1255 GCTGGGTTCCAGCGTCTGCTGAGGATCTGAAGTCTGCGGTTCTTCCGGGAGTTGACT 1314

RESULT 15
US-09-338-132-1
; Sequence 1, Application US/09338132
; Patent No. 6040164
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT FILING DATE: 1999-06-22
; EARLIER APPLICATION NUMBER: US/09/338,132
; EARLIER FILING DATE: 1999-06-22
; EARLIER APPLICATION NUMBER: 08/860,150
; EARLIER FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: PCT/EP95/05052
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: 94810746.1
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patentin Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2101
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (132)..(1499)
; US-09-338-132-1

Query Match      4.3%; Score 172.4; DB 3; Length 2101;
Best Local Similarity 49.9%; Pred. No. 4.3e-35;
Matches 539; Conservative 0; Mismatches 511; Indels 30; Gaps 3;

QY 2197 AGCGTCTACATCTATAAACAATAGAAATGAATGATGCGGGTTGGATTATCTCAAG 2256
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 262 AGCGAAGCAGCGCCCTCGCAAGCTCAGGCTCAGCTGAAGGACGAGAGCTTCGCGAGG 321
QY 2257 ATGCCAGCATCAATAGAGAGATGCTTTGCCAAAAGAACTAATACATCCGCTTAA 2316
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 322 CGCAGCGCCAGGAGAGCGCTCTGCAGCATGCCCAAGAGGAGCGAGTATCTCCGGCTGA 381
QY 2317 AAAGGCTAAAATGGCAAGTCTATTTCTGTAAGATAAAGACACTAGGAATAGGAGCAT 2376
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 382 ACGATTGCGCTCGGTGGAGGACTTTGAGGCCCTCAAGATCATCGGCGCGGT 441
QY 2377 TTGGTGAATCTGTCTAGCAAGAAATAGATAGTAAAGCTTTGTATGCAACAAAACCTC 2436
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 442 TCGTGAAGTGGCTTTGGTGCAGAAAAGGACACTGGACATGTGTGCGCCCATGAAGGTGC 501
QY 2437 TTCGAAGAAAGATGTTCTTCTTCGAAATCAAGTCTCATGTTAAGGCTGAGAG---AG 2493
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 502 TCGGCAAGCGGCATCTCTGGAAGAGGAGGAGGTGGCACACGTACCGCGGAGGCTGTCG 561
QY 2494 ATATCTGCTGAAGCTGACAATGAATGGTAGTTCTGCTCTATATTATTCATTCCAAGATA 2553
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 562 ATGCTCTGCTGAGGCGGATCATCAGTGGTGTGAAGATGTACTACAGTTTCCAGGATC 621
QY 2554 AGCAAAATTAATTTTGAATGGACTACATCTCTGGGGGTGATATGATGAGCCTATTAA 2613
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 622 CGGTCAATTTATTTGATAATAGAGTCTTTCCTGGTGGTGATATGATGACGCTTTTAA 681
QY 2614 TTAGAATGGCATCTTTCCAGAAAGCTGCGAGATTTCTACATAGACAGAACTTACCTGTG 2673
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 682 TGAAGAGCAGCGCTATCCGAGGAGGCGACACAGATTTCTATATCATGACGCGCATGG 741
QY 2674 CAGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGATATTAACCTGATAATATT 2733
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 742 CGATCATTTTATTCACAACTCGGTTTCATACACAGGATATACAGCCGATACCTTGC 801
QY 2734 TGATTGATCGTGAATGATATTAATTAATGACTGATTTGGCCCTCTGCATCGGCTTCAGAT 2793
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 802 TGTGAGCGCGGAGGCGCATCTGAAGCTCTCCGACTTCGCACTGTGCACTGGCTTAAAGA 861
QY 2794 GGACACAGATCTTAAGTACTATCAGATGGTGACCATCCAGCGGCAAGATAGCATGGATT 2853
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 862 AGTCGATCGAAGACAGACTTTTATTCGGGACTTTGTGCGAGGC-----GAAACCATCCGAT 914
QY 2854 TCAGTAATGAATGGGGGGATCCCTCAAGCTGTGCGATGTGAGACAGACTGAAGCCCATTAG 2913
    ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
```

```

Db 442 TCAGTGAAGTCGCTTTGGTCAGAAAAAGGACACTGGACATGTGTGCGCCATGAAGGTGC 501
Qy 2437 TTCGAAAGAAAGATGTTCTTCTTCGAATCAAGTCGCTCATGTTAAGCTGAGAG---AG 2493
Db 502 TGCACAAAGCGGACATGCTGGAAGAGGAGCAGAGTGGCACACGTACGGCGCGGAGGGTCTGC 561
Qy 2494 ATATCCTGGCTGAAGCTGACAAATGAATGGGTAGTTCGTCTATATATTATTCATCCAAAGATA 2553
Db 562 ATGCTCTGGTCGAGCGCGATCATCAGTGGGTGGTGAAGATGTACTACAGTTTCCAGGATC 621
Qy 2554 AGGACAAATTTATCTTGTAAATGGACTACATTCCTGGGGGTGATATGATAGCCCTATTAA 2613
Db 622 CCGTCAATTTATATTGATAATGAGTCTTGCCTGGTGGTCATATGATACGCTTTTAA 681
Qy 2614 TTAGAATGGGCATCTTCCAGAAAGTCTGGCACAGATCTACATAGCAGAACTTACCTGTG 2673
Db 682 TGAAGAAGGACACGCTATCCGAGGAGGACACAGTTCTATATCAGTGAGACGGCAFTGG 741
Qy 2674 CAGTTGAAGTGTTCATAAAATGGTATTATTCATAGAGATATTAAACCTGATATATTT 2733
Db 742 CGATCGATTCTATTACAAACTCGTTTCATACAGGGATATCAAGCCGATTAACCTGC 801
Qy 2734 TGATTTGATCGTATGATGATATTAATTAAGTACTTGGCTCTGCACTGGCTTTCAGAT 2793
Db 802 TGCTGGACGCGGAGGGCATCTGAAGCTCTCCGACTTCGGACTGTGCACTGGCTTAAAGA 861
Qy 2794 GGACACAGATTTCTAAGTACTATCAGAGTGGTGACCATCCAGCGCAAGATAGCATGGATT 2853
Db 862 AGTCGATCGAACAGACTTTTATCGGGACTTGTGCGAGC-----GAAACCATCCGAT 914
Qy 2854 TCAGTAATGAATGGGGGATCCCTCAAGCTGTCGATGTGGAGACAGACTGAAGCCATTAG 2913
Db 915 TTTATAGGCACGTGGCCAGTCCGATGGACTCCAAAGCGACGTGCGAGTCTGTGAAGCGA 974
Qy 2914 AGCGGAGAGTGCAGCGCAGCAGCAGCATGTCTAGCACATTTCTTGGTTGGGACTCCCA 2973
Db 975 AATCGACGCGCCCTCGCCTACAGCACC-----GTGGGAACGCCGG 1014
Qy 2974 ATTATATTGCACCTCAAGTGTGTACCAACAGGATACACAGTTGTGTATTGGTGGA 3033
Db 1015 ACTATATTGCAACCCGAAGTATTTCTGCAGACTGGCTACGGACCCGCTGCGACTGGTGGT 1074
Qy 3034 GTGTTGGTGTATTCTTTTGAATGTTTGGTGGACAACTCTCTTTTGGCACAAACAC 3093
Db 1075 CCCTGGGAGTCATCATGTACGAAATGCTGATGGGCTATCCTCCATTCTGCTCGGACAATC 1134
Qy 3094 CATTAAGAACACAAATGAAGTTTATCAACTGGCAACATCTTTCACATTCACACCACAAG 3153
Db 1135 CCCAGGACACCTTACCGCAAGGTGATGAATGGCGGAGACGCTGATATTTCCCCCAAGAG 1194
Qy 3154 CTAAACTCAGTCTCAAGCTTCTGATCTTATTATTAACCTTTCGAGGAGCCCGAAGATC 3213
Db 1195 ATCCCATATCGGAGGAGGCGCCAAGGAGAGCATCATCACTTCTGCTCGAGGCGCATGCC 1254
Qy 3214 GCTTAGGCAAGAATGGTCTGATGAATAAAGCTCATCCATTTTAAAAACAATGTACT 3273
Db 1255 GCTGGGTTCCAGCGCTGCTGGAGGATCTGAAGTCTGCCGTTCTTCCGGGGAGTTGACT 1314

```

Search completed: January 16, 2003, 21:55:17
 Job time : 162.058 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: January 16, 2003, 09:56:42 ; Search time 110.556 Seconds
(without alignments)
16067.254 Million cell updates/sec

Title: US-09-763-334-1
Perfect score: 3984
Sequence: 1.accttgggtgctgggacg.....agagtttttgaatttttt 3984

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	655.4	16.5	2043	9	US-09-836-392-2
2	609.8	15.3	638	10	US-09-879-536-26
3	452.8	11.4	678	10	US-09-879-536-66
4	345.2	8.7	676	9	US-09-764-868-214
5	261.4	6.6	1689	9	US-09-938-842A-1861
6	254.6	6.4	3583	9	US-09-974-298-152
7	216.8	5.4	1452	9	US-09-938-842A-2402
8	186	4.7	734	9	US-09-764-868-196
9	171.8	4.3	1818	10	US-09-771-161A-89
10	122.8	3.1	2706	12	US-10-071-751-20
11	120.6	3.0	1416	9	US-09-938-842A-2503
12	120.2	3.0	5313	10	US-09-801-368-297
13	119.2	3.0	1398	9	US-09-938-842A-633
14	117	2.9	568	9	US-09-796-692-8122
15	117	2.9	6409	10	US-09-864-864-293
16	116.2	2.9	2637	10	US-09-799-875-3
17	116.2	2.9	3003	10	US-09-799-875-1
18	112.6	2.8	2365	9	US-09-981-353-6
19	112.4	2.8	277	10	US-09-294-093B-2043

20	112.4	2.8	1244	10	US-09-771-161A-38	Sequence 38, Appl
21	112.4	2.8	1393	10	US-09-771-161A-37	Sequence 37, Appl
22	111.6	2.8	1735	9	US-09-764-868-58	Sequence 58, Appl
23	111	2.8	2311	10	US-09-810-808-6	Sequence 6, Appl
24	111	2.8	2370	9	US-10-000-039-1	Sequence 1, Appl
25	111	2.8	2370	10	US-09-969-347-214	Sequence 214, App
26	111	2.8	2370	10	US-09-880-107-3855	Sequence 3855, Ap
27	107.4	2.7	587	9	US-09-764-868-217	Sequence 217, App
28	106.4	2.7	1194	10	US-09-801-368-399	Sequence 399, App
29	105.2	2.6	1479	10	US-09-771-161A-46	Sequence 46, Appl
30	101.4	2.5	1515	10	US-09-804-471A-1	Sequence 1, Appl
31	101.4	2.5	5877	12	US-10-028-946-3	Sequence 3, Appl
32	101.4	2.5	6165	12	US-10-028-946-1	Sequence 1, Appl
33	100.4	2.5	2261	9	US-09-954-531-141	Sequence 141, App
34	100.4	2.5	2261	9	US-09-954-531-570	Sequence 570, App
35	99	2.5	1197	10	US-09-801-368-403	Sequence 403, App
36	98.2	2.5	2346	10	US-09-934-456-1137	Sequence 1137, Ap
37	96.6	2.4	6159	9	US-10-017-216-3	Sequence 3, Appl
38	96.6	2.4	6574	9	US-10-017-216-1	Sequence 1, Appl
39	95	2.4	2512	10	US-09-784-249-1	Sequence 1, Appl
40	93.4	2.3	1143	10	US-09-801-368-401	Sequence 401, App
41	93.4	2.3	2760	9	US-10-098-841-195	Sequence 195, App
42	92.6	2.3	2549	10	US-09-880-107-3691	Sequence 3691, Ap
43	92.2	2.3	1470	9	US-09-938-842A-1978	Sequence 1978, Ap
44	91.4	2.3	362	9	US-09-796-692-7789	Sequence 7789, Ap
45	91.4	2.3	3061	10	US-09-880-107-2146	Sequence 2146, Ap

ALIGNMENTS

RESULT 1
US-09-836-392-2
; Sequence 2, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-2

Query Match	16.5%	Score 655.4;	DB 9;	Length 2043;
Best Local Similarity	69.2%	Pred. NO. 1.6e-148;		
Matches '909;	Conservative 0;	Mismatches 401;	Indels 3;	Gaps 1;
OY	2241	GTGTGATTATCTCAGATGCCAGGATCAATGAGAAAGATGCTTTGTCACAAAAGAAATCT	2300	
Db	218	GCTGGACTGTGTGAAGCTGAGCAGGAGATGCGGAAGATCCTCTACGAGAAAGATCT	277	
OY	2301	AAATTACATCCGCTCTTAAAGAGGCTAAATGGAACAAGTCTATCTTTGTGAAGATAAGAC	2360	
Db	278	AATTACACAGCTTTAAAGAGGCCCAAGATGGACAGTCTATCTTTGTCAAGATCAAAACC	337	
OY	2361	CTAGGAATAGGAGCATTTGGTGAAGTCTCTCTAGCAAGAAAAGTAGACTACTAAGCTTTG	2420	
Db	338	CTGGGGATCGGTGCTCTTTTGGAGAAGTGTGCTTGTGTAAGGTGGACACTCACGCCCTG	397	

QY 2658 GCAGAACTACCTGTCAGTGAAGTG---TTTCAAAAATGGGTTTTTATTATAGAGAT 2714
DB 157 CCAGAACTACCTGTCAGTGAAGTGTTCCTTAAATGGGTTTTTATTATAGAGAT 98
QY 2715 ATTAACCTGATAATATTTTGTGATGTCGTGATGTCATATTAATTAATGACTTTGGC 2774
DB 97 ATTAACCTGATAATATTTTGTGATGTCGTGATGTCATATTAATTAATGACTTTGGC 38
QY 2775 CTCTGCACTGGCTTCAGATGGACACACGATTTCTAAGT 2811
DB 37 CTCTGCACTGGCTTCAGATGGACACACGATTTCTAAGT 1

RESULT 3
US-09-879-536-66
; Sequence 66, Application US/09879536
; Patent No. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 66
; LENGTH: 678
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc.feature
; LOCATION: (1)...(678)
; OTHER INFORMATION: n = A,T,C or G

US-09-879-536-66

Query Match 11.4%; Score 452.8; DB 10; Length 678;
Best Local Similarity 90.1%; Pred. No. 8.2e-100;
Matches 548; Conservative 0; Mismatches 50; Indels 10; Gaps 6;

QY 2180 ACTCAATCTCATCAGCAGCGCTACATCGTAAACAAATAGAGATGAATGATCG 2239
DB 1 ACTCAATCTCATCAGCAGCGCTACATCGTAAACAAATAGAGATGAATGATCG 60

QY 2240 GTTGGATATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAAAAGATC 2299
DB 61 GTTGGATATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAAAAGATC 120

QY 2300 TAATACATCTCGTTTAAAGGGCTTAAATGGACAAGTCTATGTTGTGAAGATAAGAC 2359
DB 121 TAATACATCTCGTTTAAAGGGCTTAAATGGACAAGTCTATGTTGTGAAGATAAGAC 180

QY 2360 ACTAGGAATAGGAGCAATTTGGTGAAGTCTGTAGCAAGAAAAGTAGACTAAGGCTTT 2419
DB 181 ACTAGGAATAGGAGCAATTTGGTGAAGTCTGTAGCAAGAAAAGTAGACTAAGGCTTT 240

QY 2420 GTATGCACAAAACCTCTTCAAGAAAGATGTTCTTCTTGAATCAAGTCGCTCATGT 2479
DB 241 GTATGCACAAAACCTCTTCAAGAAAGATGTTCTTCTTGAATCAAGTCGCTCATGT 300

QY 2480 TAAGCTGAGAGAGATATCCCTGGCTGAAGCTGACAAATGAATGGGTAGTCTCTATATTA 2539
DB 301 TAAGCTGAGAGAGATATCCCTGGCTGAAGCTGACAAATGAATGGGTAGTCTCTATATTA 360

QY 2540 TTCATTCCAAGATAAGGACAATTTTATCTTTGTAATGAGCTACATTTCTGGGGTGATAT 2599
DB 361 TTCATTCCAAGATAAGGCGCATTTATCTCTT--GTAATGGCTACATTTCTCTNGGGTGATAT 418

QY 2600 GATGAGCTATTAATTAAGATGGGCATC-TTTCAGAAAGTCTGGCAGATTTCTACAT-- 2656
DB 419 GAAGAGCCCATTAATTAATTAAGATGGGCATCTTTTCCAGAAAGGCTNGCACAATCTACCTTA 478

QY 2657 AGCAGAACTTACCTGTG-CAGTTGAAAGTGTTCATAAAA--TGGGTTTTTATTATCATAGAGA 2713
DB 479 GCCAGAACTTACCTGTGCGCNGTGTGAAGTGGTCTTAAATGGGGTTTAAATCTTAGAGA 538

QY 2714 TATTAAACCTGATAATATTTGATGATGATC--GTGATGTGCATATTAATTAATGACGTTT 2771
DB 539 TTTTAACTGATAATATTTGANTGGACCGGAAGGGCTTATTAAAAATGGCTTGCTTT 598

QY 2772 GGCCTCTG 2779
DB 599 GGCCTTNG 606

RESULT 4
US-09-764-868-214
; Sequence 214, Application US/09764868
; Patent No. US20020168711A1
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (628)
; OTHER INFORMATION: n equals a,t,g, or c

US-09-764-868-214

Query Match 8.7%; Score 346.2; DB 9; Length 676;
Best Local Similarity 71.2%; Pred. No. 4.7e-74;
Matches 481; Conservative 2; Mismatches 191; Indels 2; Gaps 2;

QY 2145 TTTAAATTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGCGTCTA 2204
DB 1 TTTAAATTTTATGGAGCAGCAGCTGGAGAATGTATCAAAACCTACCAGCAGAAAGTT 60

QY 2205 CATCGTAAAAACAATTAGAGAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAG 2264
DB 61 AACCGAGGCTGACGTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAG 120

QY 2265 GATCAATAGAAAGATGCTTTGCCAAAAGAATCTAATTAATACATCGTCTTAAAAGGCT 2324
DB 121 GAGCAGATGCGGAAGATCTCTACCAGAAAGATCTAATTAACACAGGTAAAGAGGGCC 180

QY 2325 AAAATGGACAAGTCTATCTTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGA 2384
DB 181 AAGATGGACAAGTCTATCTTTGTCAAGATCAAAACCTGGGATCGGCTTTTGGAGAA 240

QY 2385 GTCTGCTTAGCAAGAAAGTAGACTAAGGCTTTGTATGCACAAAACACTCTTCGAAG 2444
DB 241 GTGTCCTTGTCTAAGTGGACACTCACGCCCTGTACGCCATGAAGACCTTAAGAAA 300

QY 2445 AAGATGTTCTTCTTCGAAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTCGCT 2504

Db 301 AAGATGTCTGACCGGATCAGTGGCCCGGTCAGGCCGAGAGGACATCTCTGGCC 360
 QY 2505 GAAGCTGACAAATGAATGGGTAGTTCCTATATATTTATTTCAATTCAGAGTAAAGGACAAATTTA 2564
 Db 361 GAGCGACAGCAATCAGTGGGT-GGCAAACTCTACTACTCTCTTCCAAAGACAAAGACAGCCTG 419
 QY 2565 TACTTTGTATGACTACATTTCTGGGGTGATATGATGAGCCTATTAATAGAAATGGGC 2624
 Db 420 TACTTTGTATGACTACATTTCTGGGGTGATATGATGAGCCTATTAATAGAAATGGGC 479
 QY 2625 ATCTTTCCAGAAATGTCGGCACGATTTCTACATAGCAGAACTTACCTGTGCGAGTTGAAAGT 2684
 Db 480 GTCTTCCCTGAGCACCTGGCGGCTCTACATCGCA-RACTGACTTTGGCCATTGAGAGT 538
 QY 2685 GTTCATAAAATGGGTTTTATTCATAGAGATTAATAACCTGATATAATTTGATGATCGT 2744
 Db 539 GTCCACAAGATGGGCTTCATCCACCGAGACATCAAGCCTGATTAACATTTTGTATGATCTG 598
 QY 2745 GATGGTCATATAAATGACTGACTTTGGGCTCTGCACTGGCTTCAGATGGACACAGAT 2804
 Db 599 GATGGTCATATAAATGACTGACTTTGGGCTCTGCACTGGCTTCAGATGGAGTCAAT 658
 QY 2805 TCTAAGTACTATCAGA 2820
 Db 659 TTCCAATATTACCAGA 674

RESULT 5

US-09-938-842A-1861
 : Sequence 1861, Application US/09938842A
 : Patent No. US20020160378A1
 : GENERAL INFORMATION:
 : APPLICANT: Harper, Jeff
 : APPLICANT: Kreps, Joel
 : APPLICANT: Wang, Xun
 : APPLICANT: Zhu, Tong
 : TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 : TITLE OF INVENTION: SAME, AND METHODS OF USE
 : FILE REFERENCE: SCRIP1300-3
 : CURRENT APPLICATION NUMBER: US/09/938, 842A
 : CURRENT FILING DATE: 2001-08-24
 : PRIOR APPLICATION NUMBER: US 60/227,866
 : PRIOR FILING DATE: 2000-08-24
 : PRIOR APPLICATION NUMBER: US 60/264,647
 : PRIOR FILING DATE: 2001-01-16
 : PRIOR APPLICATION NUMBER: US 60/300,111
 : PRIOR FILING DATE: 2001-06-22
 : NUMBER OF SEQ ID NOS: 5379
 : SEQ ID NO 1861
 : LENGTH: 1689
 : TYPE: DNA
 : ORGANISM: Arabidopsis thaliana
 : US-09-938-842A-1861

Query Match 6.6%; Score 261.4; DB 9; Length 1689;
 Best Local Similarity 52.7%; Pred. No. 2.4e-53;
 Matches 642; Conservative 0; Mismatches 566; Indels 11; Gaps 3;
 QY 2142 GCATTTAAATTTTATGGGACACATGTAGAAATGTACTCAATCTCATCAGCGGT 2201
 Db 157 GCACGGAAGCAATATGCGAAGATCATATAAATTCAGAAAGAGTCTCCAGGAAGA 216
 QY 2202 CTACATCTAAAAACAATTAGAATGAAATGATCGGGTTGGATTATCTCAAGATGCC 2261
 Db 217 AAAGAGCTCGGAGCATCTTGGACAAACCTAGCTGATGCTGATCTTCTGGAAGAC 276
 QY 2262 CAGATCAAAATGAGAAAGATCGTTTGGCAAAAGAAATCAATTTACATCCGCTTTAAAGG 2321
 Db 277 AAGATGGATATTAAGAACTTTGAGAAAAAGGAAATGGAGTATATGCGTCTACAAGA 336
 QY 2322 GCTAAATGGCAACTCTATCTTCTGAGATAAGACACTAGGAATAGGAGCATTTGGT 2381

Db 337 CAGAAAAATGGGGTTTCATGACTTTGAACATCTTAAGCTGCTTAAGCTGCTTTCGGGT 396
 QY 2382 GAAGTCTGCTAGCAAGAAAGTAGATATAAGGCTTTTATATCAACAAAAAATCTTCCA 2441
 Db 397 GAGGTGAGAAATTTGTAAGAAAAATCTACTGGAAGCTATATGCAATGAAAAAGTTAAG 456
 QY 2442 AAGAAAGATGTTCTTCTTTCGAAATCAAGTCGCTCATGTTTAAAGCTGAGAGATATCCTG 2501
 Db 457 AAATCCGAGATGCTTCGACGAGGAGGCTGGAACATGTTTAAAGCTGAAAGAAATGTCTT 516
 QY 2502 GCTGAAGCTGACAAATCAATGAGGCTGCTCTATATTTATTTATTTCCCAAGTAAGACAT 2561
 Db 517 GCAGAAAGTGGATAGCCCTTTCATGTCAGCTTTGTTACTCTTTCCCAAGATGAGCAT 576
 QY 2562 TTATATCTTTGTAATGACTACATTTCTGGGGTGATATGATGAGCCTATTAATTAAGAATG 2621
 Db 577 TTGTATCTTTATGGAATACCTCCCTGAGGTGATATGATGACACTGCTGATGCGAAG 636
 QY 2622 GGATCTTTCCAGAAAGTCTGGACGATTTCTACATAGAGAACTTACCTGTGCAAGTTGA 2681
 Db 637 GATACCTTACGGGAGATGAGACTCGGTTTTATTTGTCGACAGCAATTTCTGGCTATTGAG 696
 QY 2682 AGTGTTCATAAAATGGGTTTTATTTATGATAGATATTAAACCTGATATAATTTTGTATGAT 2741
 Db 697 TCTATCCATAAGCATTAATAGTCCACAGGATATAAGCCTGATTAATTTATGATTACT 756
 QY 2742 CGTGATGGTCATATAAATTTGACTGACTTTGGCCTCTGCACTGGCTTCAGATGGACACAC 2801
 Db 757 CGAAACGGCCATATCAAGCTTTTCAGATTTTGGATTGAGCAAGTCTCTGGAAGCAAAAAT 816
 QY 2802 GATTCATAGTACTATCAGAGTGTGACCATCCAGCGCAAGATAGCATGATTTCACTAAT 2861
 Db 817 TTTCCAGATTTTAA-----GGCGAGGCTTGTGACAGAGTACAAAGCCTCGACAGACA 869
 QY 2862 GAATGGGGGATCCCTCAAGCTCTGATGTGGAGAC-AGACTGAAGCAATTTAGAGCGGAG 2920
 Db 870 ACATGATAGACTCTCAAGCCTCTCTGACCTAGAGAACTCAGGAGAACAGCTTTT 929
 QY 2921 AGCTGACCGCCAGCACAGCGATGCTTAGCACAATTTCTTTGGTGGGACTCCCAATATAT 2980
 Db 930 ACATTTGCAACAAAAACAGAAAGGACCTCGCTTTTCTACAGTAGGAATCCCGGATTACAT 989
 QY 2981 TGCACCTGAAGTGTGCTAGCAAGATACACAGTGTGATTTGGGAGTGTGG 3040
 Db 990 TGCCCTCTAGGTGCTGCTGAAGAAAGGATGGAATGGAGTGTGATTTGGTCTCTTGG 1049
 QY 3041 TGTATTTCTTTTGAATGTTGGTGACAACTCTCTTTCTTGGCACAACACCAATFAGA 3100
 Db 1050 AGCAATCATCTTCAGATGCTTGTAGGGTTTCCGCCATTTCTATTCGAAGAACCTTTGGC 1109
 QY 3101 AACACAAATGAAGTTATCACTGCGCAACATCTCTTCACATTCACCACCAAGCAACT 3160
 Db 1110 AACATGTAGAAAGATTGTAACCTGGAACCTGCTTGAAGTTCCCTGATGAAGCTAAGCT 1169
 QY 3161 CAGTCTGAAGCTCTGATCTTATTAATTAAGCTTTGCGGAGGACCGCAAGATCGCTTAGG 3220
 Db 1170 CTCCATCGAGGTAAAGATCTTATCCGAAAGCTGCTGCAATGTCGAACAGAGGCTTGG 1229
 QY 3221 CAAGAATGGTCTGATGAAATAAAGCTCATCCATTTTAAACAAATTCACITCTCCAG 3280
 Db 1230 GACCAAGAGATTCAAGAAATCAAGCAACACCTCTTGGTTTGGGGAGTCCGAATGGGAACG 1289
 QY 3281 TGACCTGAGACAGAGTCTGCTTCATACATCTCTTAATACACACCCACACATACATC 3340
 Db 1290 ---GCTATATGAGTCAAAATGCTCCATATATACCACAAGTGAAGCATGAGCTTGATACCCA 1346
 QY 3341 AAATTTTATCTCTTCAT 3359
 Db 1347 AAATTTTGAAGTTTCTGAT 1365

Db 217 GACCTGGAGCAAGAGACAGAGCTTTATGAGCTTAAAGGAACAAGATTAGTGTGAT 276
 Qy 2340 ATGTTTGTGAAGATAAGACACTAGGAATAGGACATTTGGTGAAGTCTGTCTACGAAGA 2399
 Db 277 GACTTTGAGCTTTTGAAGTATCATGTTGAGAGAGTCTTTGGTGAAGTCTGTCTACG 336
 Qy 2400 AAGTAGATCTAAGGCTTTGATGCAACAACTCTTCGAAAGAAAGATGTTCTTCTT 2459
 Db 337 GAGAAAGCTGGAATATTTATGCCATGAAGAAGTTAAGAAATCTGAATGCTCATG 396
 Qy 2460 CGAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTCGCTGAGCTGACAATGAA 2519
 Db 397 AGAGACAGGTTGAGCATGTGAGAGAGAGAGAACTGCTGCTGAGGTTGAAGCCAT 456
 Qy 2520 TGGGTAGTCTCTATATTTATTCATTCGAGATGAAGACAAATTTATACCTTTGTAATGAC 2579
 Db 457 TATATTGGAAGCTTTTACTATTCATTTTCCAGATCCCGAGTATCTGTATCTGATTGAA 516
 Qy 2580 TACATCTCGGGGTGATATGATGAGCTTATTAATTAAGATGGCATCTTTCCAGAAAGT 2639
 Db 517 TATCTCCCGGTGGTATGATGATGACCTTGTCTCATGAGGAGGACACATACGGGAAGAT 576
 Qy 2640 CTGGCAGCATTTTACATAGCAGAACTTACCTGTGCAAGTGAAGTGTTCATPAATAATGGT 2699
 Db 577 GTTCCAGATTTTATATGCTCAAAGTCTCTGCGCATTTGAATCCATACAGATACAAC 636
 Qy 2700 TTTATTCATAGATATTAACCTGATTAATTTGATGATGCTGATGCTGATGCTATTAATA 2759
 Db 637 TATATTCATAGGATATCAAACTGATAAATCTTTTGGACAAAGATGGGCACATGAA 696
 Qy 2760 TTGACTGACTTTGGCTCTGCACTGGCTTCAGATGGACACAGATTTCTAAGTACTATCAG 2819
 Db 697 CTCTCGGACTTTGGCTCTGTAAGCTCTGATTTGTAAGAAATTTACCTTCAATTCAGGAG 756
 Qy 2820 AGTGGTG--ACCATCCAGCGCAAGATAGCATGATTTTCAAGTAATTAATGCGGATCCCT 2877
 Db 757 AATAGGCGCAGGATGATGAACATATGTCAGAACCTATGATGATGATGATGATGATGCT 816
 Qy 2878 CAAGCTGTGATGGAGACAGACTGAGCCATAGAGCGAGAGCTGCA---CGCCAG 2933
 Db 817 GACACTGATACAGAGAGAGCTGGCGCAGTCCCGAGAACAACTTCAGCATTTGCGAGATG 876
 Qy 2934 CACGAGCATGTAGCAGATCTTTTGGTGGGACTCCCAATTAATTTGACCTGAAAGTG 2993
 Db 877 AATCCAGAAACTTAGCATTTTCAACTGTGGGAACACCGGACTATATGCTCTGAAAGTT 936
 Qy 2994 TTGCTACGACAGATACACAGATTTGATGATGATGATGATGATGATGATGATGATGATGAT 3053
 Db 937 TTGCTGAAGAAGGATATGGCATGATGATGATGATGATGATGATGATGATGATGATGAT 996
 Qy 3054 GAAATGTTGGTGGGCAACCTCTTTCTTGGCACAACACCATTAAGAAACACAATGAAG 3113
 Db 997 GAAATGCTGTTGGTATCTCTCTTTTATGCTGATGACCTATATCACTTCAGAGAG 1056
 Qy 3114 GTTATCAACTGGCAACATCTCTTCAATTTCCACACAAAGCTAACTCAGTCTGAAAGCT 3173
 Db 1057 ATCGTCCATTTGAGAAACCATTTGAATTTCTGAGGATGCGAAGTTTCTATCTGAGGCA 1116
 Qy 3174 TCTGATCTTATTAATCTTTGCGAGGACCCGAGATGCTTAGGCAAGAAATGTTGCT 3233
 Db 1117 AAGATCTTATCTGCAAGTGTGCTGCAAGCTTTGACCATGAGCTTTGCTGAGGAGGA 1176
 Qy 3234 GATGAATTAAGCTCATCTCTTTTAAACAAATTTGACTTCTCCAGTGACCTGAGACAG 3293
 Db 1177 GCCAGCAATCAAGGATCATCTTGGTTCAAGATGTTGTTGGGAAGAGCTCTATGAA 1236
 Qy 3294 CAGTCTGCTTACATCTTCTTAAATACACACCCACAGATACATCAAAATTTTGTATCCT 3353
 Db 1237 ATGAGGCTGCTTACAAACCAAGAGTGAACGACGAGCTAGATACACAAATTTTATGAAG 1296
 Qy 3354 GTTGATCTGATAAT 3369
 Db 3354 GTTGATCTGATAAT 3369

Db 1297 TTTGATGAAGTGAATT 1312
 RESULT 8
 US-09-764-868-196
 : Sequence 196, Application US/09764868
 : Patent No. US20020168711A1
 : GENERAL INFORMATION:
 : APPLICANT: Rosen et al.
 : TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 : FILE REFERENCE: PIZ32
 : CURRENT APPLICATION NUMBER: US/09/764,868
 : CURRENT FILING DATE: 2001-01-17
 : Prior application data removed - refer to PALM or file wrapper
 : NUMBER OF SEQ ID NOS: 1510
 : SOFTWARE: PatentIn Ver. 2.0
 : SEQ ID NO 196
 : LENGTH: 734
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-764-868-196

Query Match 4.7%; Score 186; DB 9; Length 734;
 Best Local Similarity 55.6%; Pred. No. 2.5e-35;
 Matches 419; Conservative 2; Mismatches 297; Indels 36; Gaps 2;
 Qy 2400 AAGTAGATCTAAGGCTTTGATGCAACAACTCTTCGAAAGAAAGATGTTCTTCTT 2459
 Db 12 AAGAAAGATACAGCCCATATCTATGCAATGAAGATATTGAGAAAGTCTGATATGCTTGA 71
 Qy 2460 CGAATCAAGTCGCTCATGTTAAGGCTGAGAGAGATATCTCGCTGAGCTGACAATGAA 2519
 Db 72 AAGAGCAGGTTGGCCCATATCCGAGCAGAAAGAGATATTTGGTARAAGCAGATGGTGCC 131
 Qy 2520 TGGTGTGCTCTATATTTATTCATTCAGATGAAGACAAATTTATCTTTGTAATGAC 2579
 Db 132 TGGTGTGAGATGTTTTCAGATTTTCAGATGAAGAGAACTTTATCTTANTCATGAA 191
 Qy 2580 TACATCTCGGGGTGATGATGAGCTTATTAATTAAGATGGCCTCTTCCAGAAAGT 2639
 Db 192 TTTCTCCCTGGAGTGACATGACATGCTAATGAAGAAAGACWCCTTGACAGAAAG 251
 Qy 2640 CTGCGCATTTACATAGCAGAACTTACCTGTGCGAGTTGAAAGTGTTCATAAATGGT 2699
 Db 252 GAAACACAGTTTCTACATTTTACAGACTGTTCTGCAATAGATGCGATCCACAGTTGGGT 311
 Qy 2700 TTTATTCATACAGATATTAACCTGATAATTTTGTGATGATGCTGATGCTATATAA 2759
 Db 312 TTATTCATCGGATATTAAGCAGACACACCTTTTATGATGCCAGGGTCAATGAA 371
 Qy 2760 TTGACTGACTTTGGCTCTGCACTGGCTTCAGATGGACACAGATTTAAGTACTATCAG 2819
 Db 372 TTATCTGATTTGGTTTATGTACGGGATTAAGAAAGCTCACAGGACTGAATTTTATAGA 431
 Qy 2820 AGTGGTGACCATCCAGCAGATAGCATGATTTCACTAATGAATGGGGGATCCCTCA 2879
 Db 432 AATCTCACACACACCCCAAGT-----G 456
 Qy 2880 AGCTGTGATGAGCAGACACTGAAGCATTAGAGGGAGAGCTGCGACGCCACCCAG 2939
 Db 457 ACTTCTCATTTCAAGACATGACTCAAGAGAAAGCAGAACTTGGAG--AAGAACAGG 515
 Qy 2940 CGATGCTAGCAGATTTTGGTGGGACTCCCAATTTATTTGACCTGAGTGTGCTA 2999
 Db 516 AGACAACCTGGCATATTCACAGTTGGGACACAGATTTACATTTGCTCCAGAAATTTATG 575
 Qy 3000 CGAAGATGACACAGTCTGTTGATTTGGTGGAGTGTGTTGTTTATTTTGAATG 3059
 Db 576 CAGACTGTTTACACAAATTTGTGACTTGGTGGTCTTTGGGAGTCATATGATGAAATG 635
 Qy 3060 TTGTTGGGACAACTCTTCTTGGCACAACACCATTTTGAAGACACAATGAAGTTATC 3119
 Db 636 CTAATAGGATATCCACCTTTCTGCTCTGAACACCTCAAGAGACATACAGAAAGTATG 695

Fri Jan 17 11:16:55 2003

Db 215 GCATTTTGGGAAGACGTCATATAATGCTCATGCAAAATCAGAAATGATGTACAA 274
QY 2532 CTATATTTATTCATCCAGACAAGCAATTTATACCTTTGTAATGGACTACATTCCTGGG 2591
Db 275 TTACATTTTGTCTTTCAGAGATCAAAATATCTTTATATGTCATGGATATATATCCGGGG 334
QY 2592 GGTGATATGATGAGCCTTAATTAAGATGGGCATCTTTCCAGAAAGTCGTGGCAGATTC 2651
Db 335 GGTGACTGGTGAGTCCTATGTCGATATGAAG---TTCCAGAAAAATGGCAATGTC 391
QY 2652 TACATAGCAGAACTTACCTGTGCAAGTGTGTAATAATGGCTTTTATTCATAGA 2711
Db 392 TATACAAATGGAAGTGTGTCATCACTTCACTCCATGGGATTTGTACATCGT 451
QY 2712 GATATTAACCTGATAATATTTGATGATGTCATATATTAATGACTGACTTT 2771
Db 452 GATGTTAAACCTGATAATATGCTTACAGAAATATGCTATTTAAAGTTAGCTGACTTT 511
QY 2772 GGCCTCTGCA 2781
Db 512 GGAACCTGTA 521

RESULT 11
US-09-938-842A-2503
; Sequence 2503, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SCRI1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR FILING DATE: 2000-08-24
; PRIOR FILING DATE: 2000-08-24
; PRIOR FILING DATE: 2001-01-16
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2503
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2503

Query Match: 3.0%; Score 120.6; DB 9; Length 1416;
Best Local Similarity 55.0%; Pred. No. 2.3e-19;
Matches 237; Conservative 0; Mismatches 194; Indels 0; Gaps 0;
QY 2343 TTTGTGAAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTAGCAAGAAA 2402
Db 418 TTTGAGTATTGAAGTTTGGGAGAGGTGCAATTTGGAAGAGTGTACCAAGTCAAGAAA 477
QY 2403 GTAGATTAAGGCTTTGTATGCAACAAACCTTCGAAAGAAAGATGTTCTTCGA 2462
Db 478 AAAGACAGCTGTGATATACCGCATGAGGTCTATGAGAAAGATAAAATGTTGAGAAG 537
QY 2463 AATCAAGTCGCTATGTTAAGGCTGAGAGATATCTCGCTGAAGCTGCAATGAATGG 2522
Db 538 AATCATGCTGAATCATGCAAGCGGCGGATATTTACCAAAATCGCATCTCTTC 597
QY 2523 GTAGTTCGCTATATTTATTCATTCAGATAAGCAATTTATACCTTTGTAATGGACTAC 2582
Db 598 ATTGTGCAACTTAATACTCTTTTCCAGACCAATACAGATGATCTGTTCTTCTGCTTT 657
QY 2583 ATTCTGGGGTGTATCATGAGGCTTATTAATAGAAATGGCACTCTTCCAGAAAGCTG 2642

Db 658 ATAAACGAGGTCATCTTTTCTCCAGCTCTATCACCAGGCGCTTTTCAGGAGGACTTG 717
QY 2643 GCACGATTTACATACAGCAAACTTACCTGTGAGTGAAGTGTCAAAAATGGTGT 2702
Db 718 GCTCGTGTGTACACTCAGAAATCGTCTCTGAGTGTTCCTCCATCTCCATGAGAAAGCATA 777
QY 2703 ATTCTATAGAGATTAATAACCTGATAATATTTGATGATCGTGTGATGTCATATTAATG 2762
Db 778 ATCCATAGAGATCTTAACCTGAAACATACATCATGAGCTAGATGGCCATCTGATGTTA 837
QY 2763 ACTGACTTTGG 2773
Db 838 ACTGATTTGG 848

RESULT 12
US-09-801-368-297
; Sequence 297, Application US/09801368
; Patent No. US20020128250A1
; GENERAL INFORMATION:
; APPLICANT: Busby, Robert
; APPLICANT: Cali, Brian
; APPLICANT: Hecht, Peter
; APPLICANT: Holtzman, Doug
; APPLICANT: Madden, Kevin
; APPLICANT: Maxon, Mary
; APPLICANT: Milne, Todd
; APPLICANT: No. US20020128250A1man, Thea
; APPLICANT: Royer, John
; APPLICANT: Salama, Sofie
; APPLICANT: Sherman, Amir
; APPLICANT: Silva, Jeff
; APPLICANT: Summers, Eric
; TITLE OF INVENTION: Methods for Improving Secondary Metabolite Production in Fungi
; FILE REFERENCE: 109272.147
; CURRENT APPLICATION NUMBER: US/09/801,368
; CURRENT FILING DATE: 2001-03-07
; PRIOR APPLICATION NUMBER: US 09/487,558
; PRIOR FILING DATE: 2000-01-19
; PRIOR FILING DATE: 1999-10-20
; NUMBER OF SEQ ID NOS: 440
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 297
; LENGTH: 5313
; TYPE: DNA
; ORGANISM: Saccharomyces cerevisiae
US-09-801-368-297

Query Match: 3.0%; Score 120.2; DB 10; Length 5313;
Best Local Similarity 56.6%; Pred. No. 5.8e-19;
Matches 243; Conservative 0; Mismatches 183; Indels 3; Gaps 1;
QY 2353 TAAAGACACTAGGAATAGAGAGCATTTGGTGAAGTCTGTAGCAAGAAAGTACTACTA 2412
Db 2390 TGAACCAATACAGAAAGTGTATGTTAGTGTATCTAGCAGGAAAGAACTACAG 2449
QY 2413 AGCTTTGTATCCACAAACACTCTCGAAAGAAAGATGTTCTCTTCCAAATCAAGTCG 2472
Db 2450 GAGATATTTTGTCTATAAGGTTCTAAGGAATCAGATATGATGCCCAAAATCAAGTAA 2509
QY 2473 CTCATCTTAAGGCTGAGAGAGATATC---CTGCTGAAGCTGACAATGAATGGTAGTTC 2529
Db 2510 CAATGCTCAATCCGAGAGAGCAATCATGTTTCAAGTATAGCCCTATGTTGGA 2569
QY 2530 GTCTATATTTATTCATCCAGATAGGCAATTTACTTTCTTAATGACTACTACTCTG 2589
Db 2570 GACTATTTCTAGTTTCCAAAATAAAGATAACCTTTTCTTAGTGAGTATTTACCAG 2629
QY 2590 GGGTCTGATATGATGAGCCTTATTAATTAGAATGGCATCTTTCCAGAAAGCTTGGCAGAT 2649
Db 2630 GTGAGATTTGGCCACTTTAATCAAGATGATGGGATCTCTGCCGATCAATGGCCAAGC 2689

QY 2653 ACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTTATTTCATAGAG 2712
DB 1035 AACTCTGCTGAAGTTGTTCTTGTCTGCAATACACTCCATGGGTTTAAATACACAGAG 1094
QY 2713 ATATTAACTGTATAATATTTTGAATGATCGTGCATGTCATATTAATTTGACTGCTTTG 2772
DB 1095 ATGTGAAGCTGACAAACATGCTCTTGGATAAAACATGGACATCTAAAAATTAGCAGATTTG 1154
QY 2773 GCCTCTGCA 2781
DB 1155 GCACGTGTA 1163

Search completed: January 16, 2003, 22:02:57
Job time : 183.556 secs

DB 206 TTTTATGCTTCAAGATGATAGGTATCTGTACATGGTAATGGAGTACATGCTGGTG 265
QY 2593 GTGATATGATGAGCTATTAATTAGATGGCATCTTTCCAGAAAGTCTGGCAGATTC 2652
DB 266 GAGACCTTGTAAACC---TTATAGATTAATTATGATGCTGAAATAATGGCCAAATTT 322
QY 2653 ACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTTATTTCATAGAG 2712
DB 323 AACTGCTGAAGTTGTTCTTGTCTGCAATACACTCCATGGGTTTAAATACACAGAG 382
QY 2713 ATATTAACTGTATAATTTTGAATGATCGTGCATGTCATATTAATTTGACTGCTTTG 2772
DB 383 ATGTGAAGCTTGACAAACATGCTCTTGGATAAACATGGACATCTAAAAATTAGCAGATTTG 442
QY 2773 GCCTCTGCA 2781
DB 443 GCACGTGTA 451

RESULT 15
US-09-864-864-293
; Sequence 293; Application US/09864864
; Patent No. US20020102679A1
; GENERAL INFORMATION:
; APPLICANT: Xu, Jiangchun
; APPLICANT: Mitcham, Jennifer L.
; APPLICANT: Harlocker, Susan L.
; APPLICANT: Dillon, Davin C.
; APPLICANT: Secrist, Heather
; APPLICANT: Lodes, Michael J.
; APPLICANT: Algate, Paul A.
; APPLICANT: Fling, Steve P.
; APPLICANT: Mannion, Jane
; APPLICANT: Benson, Darin R.
; APPLICANT: Carter, Darick
; TITLE OF INVENTION: COMPOSITIONS AND METHODS FOR THE THERAPY
; TITLE OF INVENTION: AND DIAGNOSIS OF OVARIAN CANCER
; FILE REFERENCE: 210121.523
; CURRENT APPLICATION NUMBER: US/09/864.864
; CURRENT FILING DATE: 2001-05-23
; NUMBER OF SEQ ID NOS: 341
; SOFTWARE: Corixa Invention Disclosure Database
; SEQ ID NO 293
; LENGTH: 6409
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-864-864-293

Query Match 2.9%; Score 117; DB 10; Length 6409;
Best Local Similarity 56.2%; Pred No. 3.8e-18;
Matches 241; Conservative 0; Mismatches 185; Indels 3; Gaps 1;
QY 2353 TAAAGACACTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATACTA 2412
DB 738 TAAAGTTTATGGAGAGGTGCTTTTGGTGAAGTGCAGTTGCTGCACAGGCATCGC 797
QY 2413 AGCCTTTGTATGCACACAAACTCTTCGAAGAAAGATGTTCTTCTCGAAATCAAGTCG 2472
DB 798 AGAAGGTTTATGCTATGAAGCTCTTAGTAAGTTTGAATGATATAAAGATCAGATTCG 857
QY 2473 CTCATGTTAAGGCTGAGAGATATCCTGGCTGAAGCTGCACAAATGAATGGTGTGCTC 2532
DB 858 CCTTTTTTTGGGAAGAAAGACATATTATGSCCTTTTGCCCAATAGCCCTGGGTGTTGAGC 917
QY 2533 TATATTATTCATCCAAAGATAGGACAATTTATCTTTGTAATGGACTACATTCCTGGGG 2592
DB 918 TTTTATGCTTTCAGATGATAGGTATCTGTACATGTAATGGAGTACATGCTGCTGGTG 977
QY 2593 GTCATATGATGACCTTATTAATTTAGAAATGGGCATCTTTCCAGAAAGTCTGGCAGCATCT 2652
DB 978 GAGACCTTGTAAACC---TTATGAGTAAATATGATGCTGCTGNAATAATGGCCARATTT 1034

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:55:17 ; Search time 3443.86 Seconds
(without alignments)
18735.612 Million cell updates/sec

Title: US-09-763-334-1

Perfect score: 3984

Sequence: 1 accttgggttgctggacg.....agagttttgaaatttttt 3984

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0.

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

```

1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estnu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_estl:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

```

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB ID	Description
C 1	930.8	23.4	1807	11 AF119846	AF119846 Homo sapi
C 2	804.8	20.2	1216	11 BC015665	BC015665 Homo sapi
C 3	697.4	17.5	721	9 A1114863	A1114863 HAL1474 Hu
C 4	650.8	16.3	694	10 AV708169	AV708169 AV708169
C 5	590.6	14.8	657	14 BM971435	BM971435 UI-CF-D01
C 6	560	14.1	1203	12 BG331628	BG331628 602432628

C 7	559.8	14.1	637	9 AA910802	AA910802 ol48d06.s
8	555.2	13.9	578	10 BE093083	BE093083 RC5-BT074
9	548.2	13.8	597	14 BQ309028	BQ309028 MKO-BT450
10	543.2	13.6	550	10 AW502306	AW502306 UI-HF-BRO
11	533.4	13.4	855	10 BE568367	BE568367 601341939
12	530.6	13.3	673	10 BE557112	BE557112 BB557112
13	518.4	13.0	648	13 BI155504	BI155504 602903902
14	511	12.8	658	12 BG088986	BG088986 ut60g11.y
C 15	497.6	12.5	665	17 AZ340286	AZ340286 IM0072J17
16	495.2	12.4	555	12 BE723541	BE723541 193387 MA
17	491.6	12.3	667	10 BE174219	BE174219 BE174219
18	474.2	11.9	858	12 BE973820	BE973820 601680625
19	466.8	11.7	488	10 BE093084	BE093084 RC5-BT074
20	449.8	11.3	453	10 BE093054	BE093054 RC5-BT074
C 21	439.6	11.0	457	13 BI034906	BI034906 OV2-NN200
22	436.4	11.0	650	13 BI930346	BI930346 603344592
C 23	432.6	10.9	474	13 BG958580	BG958580 PM2-CT026
24	424	10.6	1079	13 BM460659	BM460659 AGENCOURT
C 25	414.4	10.4	600	12 BG081357	BG081357 H3064C02-
C 26	414.4	10.4	602	14 BQ201853	BQ201853 UI-R-D01-
C 27	414.4	10.4	621	12 BG081236	BG081236 H3062G12-
28	403.2	10.1	571	10 BE395243	BE395243 601311110
29	401.4	10.1	496	9 AI594372	AI594372 VO15F09.Y
30	387.4	9.7	930	14 BQ989648	BQ989648 AGENCOURT
31	384.6	9.7	405	10 BE093082	BE093082 RC5-BT074
32	381.8	9.6	403	10 BE093081	BE093081 RC5-BT074
33	375.4	9.4	737	14 BM950668	BM950668 UI-M-EH04
C 34	367	9.2	376	10 BE088344	BE088344 CM1-BT068
35	364.4	9.1	390	14 BQ312644	BQ312644 MKO-BN011
36	361.8	9.1	391	14 BQ291582	BQ291582 PM3-AN002
37	359	9.0	559	10 AW642458	AW642458 CM19A02.W
38	349.8	8.8	473	9 AI430221	AI430221 me57h09.y
39	346.2	8.7	425	10 BB751204	BB751204 BB751204
C 40	341.4	8.6	458	9 AI596344	AI596344 me57h09.x
41	333.4	8.4	689	13 BI331257	BI331257 602982912
42	332	8.3	407	9 AA592233	AA592233 VO15F09.F
43	326.6	8.2	863	10 BE620135	BE620135 601483053
44	324.2	8.1	570	12 BF396283	BF396283 UI-R-BS2-
45	324.2	8.1	875	13 BI664597	BI664597 603290249

ALIGNMENTS

RESULT 1
AF119846/C
LOCUS AF119846 1807 bp mRNA linear HTC 08-MAY-2001
DEFINITION Homo sapiens PRO1474 mRNA, complete cds.
ACCESSION AF119846
VERSION AF119846.1 GI:7770128
KEYWORDS HTC.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE 1 (bases 1 to 1807)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.
TITLE Functional prediction of the coding sequences of 79 new genes deduced by analysis of cDNA clones from human fetal liver
JOURNAL Unpublished
REFERENCE 2 (bases 1 to 1807)
AUTHORS Zhang,C., Yu,Y., Zhang,S., Wei,H., Zhang,Y., Zhou,G., Bi,J., Liu,M. and He,F.
TITLE Direct Submission
JOURNAL Submitted (13-JAN-1999), Department of Experimental Hematology, Institute of Radiation Medicine, Beijing Taiping Road 27, Beijing, Beijing 100850, P. R. China
FEATURES
source
Location/Qualifiers
1..1807
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="FLB5626"

```

/issue_type="liver"
/dev_stage="fetus"
complement(841..1530)
/notes="predicted protein of HQ1474"
/codon_start=1
/product="PRO1474"
/protein_id="AAF69600.1"
/db_xref="GI:770129"
/translation="MAKLSRDRVFKASYTFOEVHVFIOGLPEGLMAGLQVPEVQNL
SDWFFLMCLVDHALAGSEVQVQEGLODGLKGLKGVWELVOLQDSQVGFQI
IGFGLRLYLDALQGVLRIVPVDTLEQLAHVSHLLKHRHDFCRIPQGLKAKWR
RRWPVGVKGRASRLRPLGRAALARGSPVRLPALPHOGYLGILLGRTLAASALRPVA
AVQELCSPLQS"
BASE COUNT      452 a   440 c   451 g   464 t
ORIGIN
Query Match      23.4%; Score 930.8; DB 11; Length 1807;
Best Local Similarity 99.8%; Pred. No. 4.1e-228;
Matches 932; Conservative 0; Mismatches 2; Indels 0; Gaps 0;
QY 3 CTTTGGTTGCTGGAGGACTCTGGCCGCTCAGGCTCGCCCTCAGGCCGCTGCGCCG 62
DB 934 CTTGGGTTGCTGGAGGACTCTGGCCGCTCAGGCTCGCCCTCAGGCCGCTGCGCCG 875
QY 63 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTATTTATTTAAAGAAATTTTAA 122
DB 874 TGTCCAGGAGCTCTGCTCTCCCTCCAGAGTTAAATTATTTATTTAAAGAAATTTTAA 815
QY 123 AGTCTGGGAGCTCTCTTGAAGGATCATTTTCACCTTTTCTCAGAGAAAGCTCTGGATC 182
DB 814 AGTCTGGGAGCTCTCTTGAAGGATCATTTTCACCTTTTCTCAGAGAAAGCTCTGGATC 755
QY 183 TATCAATAAAGAGCTCTCTGCTGGCTACATATATAGATGTTTTCATGAGAGGAGT 242
DB 754 TATCAATAAAGAGCTCTCTGCTGGCTACATATATAGATGTTTTCATGAGAGGAGT 695
QY 243 GAAAGCCAGAGGATATAGACAAATAGAGCCTTAAGCCTTTCCTGCCAGTAATATCT 302
DB 694 GAAAGCCAGAGGATATAGACAAATAGAGCCTTTCCTGCCAGTAATATCT 635
QY 303 GTCAGTAGCGGCAATGTTTACAAGAAATTCGGGAATCCTTAGGAATTTATCTAAACCA 362
DB 634 GTCAGTAGCGGCAATGTTTACAAGAAATTCGGGAATCCTTAGGAATTTATCTAAACCA 575
QY 363 TCTGATGCTGCTAAGGCTCAGATACATGAGTAAATGCTCAACGAGATCTCTGCACAA 422
DB 574 TCTGATGCTGCTAAGGCTCAGATACATGAGTAAATGCTCAACGAGATCTCTGCACAA 515
QY 423 GTCAGAAATCCACCAATTTGGGAGCATCATAAAGCCTTCAGGAAATTCGAACTCT 482
DB 514 GTCAGAAATCCACCAATTTGGGAGCATCATAAAGCCTTCAGGAAATTCGAACTCT 455
QY 483 CTGCTTCCATTGCAATGAACAAATTTCTCTCGAGTACTTCAGAGTTAATCCACAA 542
DB 454 CTGCTTCCATTGCAATGAACAAATTTCTCTCGAGTACTTCAGAGTTAATCCACAA 395
QY 543 ATGCTTCAAGACTTCAAGCTGCTGATTTGATGAGGATGTTATACAAAGCTCTTCAG 602
DB 394 ATGCTTCAAGACTTCAAGCTGCTGATTTGATGAGGATGTTATACAAAGCTCTTCAG 335
QY 603 AAACTTAACAACAAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGATTTACCAA 662
DB 334 AAACTTAACAACAAGATATAGAGCAGCAATTTGAATTCATTAGTAAATGATTTACCAA 275
QY 663 GATCCTCGAGGAGAGGATGCTCGAGGAGTGGCCAGAGCTTATTAATCCAGCATGAAA 722
DB 274 GATCCTCGAGGAGAGGATGCTCGAGGAGTGGCCAGAGCTTATTAATCCAGCATGAAA 215
QY 723 CCAGGAATGTGAGCAATCAGTTAAACCAACAGAGCTGGAAGGTTCTTAAGAAATCC 782
DB 214 CCAGGAATGTGAGCAATCAGTTAAACCAACAGAGCTGGAAGGTTCTTAAGAAATCC 155
QY 783 TTAGTTCTCAGAGGATGCGCCGCTCAGGAGAAAGTGTGCGCTATCATTTCTGAGAGT 842

```

```

DB 154 TTAGTTCTCAGAGGACTGGCCGCTCAGGAGAAAGTGTGGCCCTATCATTTCTGAGAGT 95
QY 843 CCCAACTCAGACAGAGTGTAGGAGAGCTTTGCTGATCTGATGATATATCAGCATTTGTT 902
DB 94 CCCAACTCAGACAGAGTGTAGGAGAGCTTTGCTGATCTGATGATATATCAGCATTTGTT 35
QY 903 CAAGCTCACCCTAGCAACGAGAGAGTGAACC 936
DB 34 CAAGCTCACCCTAGCAACGAGAGAGTGAACC 1
RESULT 2
LOCUS      BC015665      1216 bp      mRNA      linear      HTC 09-OCT-2001
DEFINITION Homo sapiens, Similar to LATS (large tumor suppressor, Drosophila)
             homolog 1, clone IMAGE:4550430, mRNA.
ACCESSION   BC015665
VERSION     BC015665.1 GI:15990533
KEYWORDS    HTC.
SOURCE      Homo sapiens.
ORGANISM    Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.
REFERENCE   1 (bases 1 to 1216)
AUTHORS    Strausberg, R.
TITLE      Direct Submission
JOURNAL     Submitted (04-OCT-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
REMARK      NIH-MGC Project URL: http://mgc.nci.nih.gov
COMMENT      Contact: MGC help desk
             Email: cgapbs-remail.nih.gov
             Tissue Procurement: DCTD/DPF/Gazdar
             cDNA Library Preparation: Rubin Laboratory
             DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
             DNA Sequencing by: National Institutes of Health Intramural
             Sequencing Center (NISC),
             Gaithersburg, Maryland.
             Web site: http://www.nisc.nih.gov/
             Contact: nisc.mgc@hgrl.nih.gov
             Shenchenko, Y., Wetherby, K.D., Beckstrom-Sternberg, S.M.,
             Benjamin, B., Blakesley, R.W., Bouffard, G.G., Brinkley, C., Brooks, S.,
             Dietrich, N.L., Guan, X., Gupta, J., Ho, S.-L., Karlins, E., Legaspi, R.,
             Lim, M., Maduro, Q.L., Masiello, C., Mastrian, S.D., McCloskey, J.C.,
             McDowell, J., Pearson, R., Snyder, B., Stantripop, S., Thomas, P.J.,
             Tjongson, E.E., Touchman, J.W., Tsurgeon, C., Vogt, J.L., Walker, M.A.,
             Zhang, L.-H. and Green, E.D.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov
Series: IRAL Plate: 33 Row: b Column: 11
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 10862687
This clone has the following problem: incomplete processing.
FEATURES
             Location/Qualifiers
             1..1216
             /organism="Homo sapiens"
             /db_xref="taxon:9606"
             /clone="IMAGE:4550430"
             /tissue_type="Lung, large cell carcinoma"
             /clone_lib="NIH_MGC_18"
             /lab_host="DH10B-R"
             /note="Vector: pOTB7"
BASE COUNT. 415 a 270 c 251 g 280 t
ORIGIN
Query Match      20.2%; Score 804.8; DB 11; Length 1216;
Best Local Similarity 90.4%; Pred. No. 9.7e-196;
Matches 911; Conservative 0; Mismatches 2; Indels 95; Gaps 1;
QY 3 CTTTGGTTGCTGGAGGAGCTCTGGCCGCTCAGGCTCGCCCTCAGGCCGCTGCGCCG 62

```

```

Db 98 CTTGGGTTGCTGGGAGGACTCTGGCCCTCAGGCTCCGCTCAGGCCGCTGGCCGC 157
QY 63 TCTCCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTGTAAGAATTTTAA 122
Db 158 TCTCCAGGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTGTAAGAATTTTAA 217
QY 123 ACTCTGGGAGCTCTCTTGAAGGATCATTTTCACATTTTCTCAGAGAAGCTCTGGATC 182
Db 218 ACTCTGGGAGCTCTCTTGAAGGATCATTTTCACATTTTCTCAGAGAAGCTCTGGATC 277
QY 183 TATCAATAAAGAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGAGT 242
Db 278 TATCAATAAAGAGTCTCTGCTGGGCTACATATATAGATGTTTTCATGAAGAGGAGT 337
QY 243 GAAAAGCCAGAGGATATAGACAAATGAGCCTAAGCCTTTCTCCCTCAGTAACTATCT 302
Db 338 GAAAAGCCAGAGGATATAGACAAATGAGCCTAAGCCTTTCTCCCTCAGTAACTATCT 397
QY 303 GTCACTAGCCGGCAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCA 362
Db 398 GTCACTAGCCGGCAATGTTTACAAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCA 457
QY 363 TCTGATGCTGCTAAGCTGAGCATACATGAGTAAATGTTCAACCGAAGATCCTCGACAA 422
Db 458 TCTGATGCTGCTAAGCTGAGCATACATGAGTAAATGTTCAACCGAAGATCCTCGACAA 517
QY 423 GTCAGAAATCCACCAATTTGGGAGGATCATATAAGCCTTCAGAGAAATTCGAATCT 482
Db 518 GTCAGAAATCCACCAATTTGGGAGGATCATATAAGCCTTCAGAGAAATTCGAATCT 577
QY 483 CTGCTTCCATTTGCAAAATGTTTCTCTCGAGTACTTCAAGATTTAATCCACAA 542
Db 578 CTGCTTCCATTTGCAAAATGTTTCTCTCGAGTACTTCAAGATTTAATCCACAA 637
QY 543 ATGCTTCAAGACTTCAAGCTCTGGATTTGATGAGT----- 578
Db 638 ATGCTTCAAGACTTCAAGCTCTGGATTTGATGAGGAGTATCTGTGCTAGTCTGT 697
QY 579 ----- 578
Db 698 TCACCCATAGTCTTACTAAGCCATTTCTCATTTTAAATGTCGTTGAAATCTGAATC 757
QY 579 -----GATATGTTTATACAGCTCTTTCAGAAACTAACACAGAGTATAGAG 627
Db 758 CTCAGGAGCAAGATATGTTTATACAGCTCTTTCAGAAACTAACACAGAGTATAGAG 817
QY 628 CAGCAATGAAATTCATTAAGTAAATGAGTTACCAAGATCCTCGAGAGCAGATGGGTG 687
Db 818 CAGCAATGAAATTCATTAAGTAAATGAGTTACCAAGATCCTCGAGAGCAGATGGGTG 877
QY 688 CAGCAGCTGCCAGACTTATTAATGCCAGCATGAAACCCAGGGAATGTCAGCAATCAGTTA 747
Db 878 CAGCAGCTGCCAGACTTATTAATGCCAGCATGAAACCCAGGGAATGTCAGCAATCAGTTA 937
QY 748 ACCGCAACAGAGCTGGAAGGTTCTTAAGAAATCTTCTAGTCTCCTCAGAGGATGCCCGC 807
Db 938 ACCGCAACAGAGCTGGAAGGTTCTTAAGAAATCTTCTAGTCTCCTCAGAGGATGCCCGC 997
QY 808 CACTAGGAGAAAGTGTGGCTATCATTTCTGAGAGTCCCAACTTCACAGCAGATGTAGAA 867
Db 998 CACTAGGAGAAAGTGTGGCTATCATTTCTGAGAGTCCCAACTTCACAGCAGATGTAGAA 1057
QY 868 GACCTTTCTGTGATCTGTATATCAGCATTTGTTCAAGCTCACCTA 915
Db 1058 GACCTTTCTGTGATCTGTATATCAGCATTTGTTCAAGCTCACCTA 1105

```

RESULT 3
 A1114863/c
 LOCUS
 DEFINITION HA1474 Human fetal liver cDNA library Homo sapiens cDNA, mRNA
 721 bp mRNA linear EST 11-NOV-1999

```

ACCESSION A1114863
VERSION A1114863.1 GI:6360208
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homidae; Homo.
YU, Y., Zhang, C., Luo, L., Ouyang, S., Li, W., Wu, J., Zhou, S., Liu, M.
and He, F.
Expression profile analysis of a human fetal liver cDNA library
Unpublished (1998)
Contact: Yongtao Yu
Department of Hematology
Beijing Institute of Radiation Medicine
27 Taiping Road, Beijing 100850, P.R.China
Tel: 0086-10-68159479
Fax: 0086-10-68214653
Email: YUT48@yahoo.com.
FEATURES
Source
1. 721
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone_lib="Human fetal liver cDNA library"
/tissue_type="liver"
/dev_stage="fetal"
/lab_host="MC1061/P3"
/note="vector: pCDNA1"
BASE COUNT 170 a 156 c 151 g 242 t 2 others
ORIGIN
Query Match 17.5%; Score 697.4; DB 9; Length 721;
Best Local Similarity 98.8%; Pred. No. 3.6e-168;
Matches 712; Conservative 0; Mismatches 8; Indels 1; Gaps 1;
QY 217 ATATAGATGTTTTCATGAGGAGGAGTGAAGAGCCAGAGGATATAGACAAATGAGCCCTA 276
Db 721 ATATAGATGTTTTCATGAGGAGGAGTGAAGAGCCAGAGGATATAGACAAATGAGCCCTA 662
QY 277 AGACCTTTCCCTG-CCAGTAACATATCTGTCAGTACCGGCAATGTTACAGAAATTCGG 335
Db 661 AGACCTTTCCCTGCCAGTAATTTATCTGTCAGTACCGGCAATGTTACAGAAATTCGG 602
QY 336 GAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGTAAGGCTGAGCATACATGAGT 395
Db 601 GAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGTAAGGCTGAGCATACATGAGT 542
QY 396 AAAATGTCACCGAAGATCCTCGACAAGTCAAGAAATTCACCCAAATTTGGGACGATCAT 455
Db 541 AAAATGTCACCGAAGATCCTCGACAAGTCAAGAAATTCACCCAAATTTGGGACGATCAT 482
QY 456 AAAGCCTTGCAGGAAATTCGAAACCTCTGCTTCCATTTGCAATGCAACAATTTCTTCT 515
Db 481 AAAGCCTTGCAGGAAATTCGAAACCTCTGCTTCCATTTGCAATGCAACAATTTCTTCT 422
QY 516 CGGAGTACTTCAAGAGTTAATCCACAAATGCTTCAAGACTTCAAGCTGCTGATTTGAT 575
Db 421 CGGAGTACTTCAAGAGTTAATCCACAAATGCTTCAAGACTTCAAGCTGCTGATTTGAT 362
QY 576 GAGGATATGGTTATACAGCTCTTCCAGAAACTAACACAGAGTATAGAGCAGCAAT 635
Db 361 GAGGATATGGTTATACAGCTCTTCCAGAAACTAACACAGAGTATAGAGCAGCAAT 302
QY 636 GAATTCATTTAGTAAATGAGTTACCAAGATCCTCGAGAGGAGATGGCTCGACGAGT 695
Db 301 GAATTCATTTAGTAAATGAGTTACCAAGATCCTCGAGAGGAGATGGCTCGACGAGT 242
QY 696 GCCAGACCTATTATGCCAGCATGAAACCCAGGGAATGTCAGCAATTCAGTAAACGCAAA 755
Db 241 GCCAGACCTATTATGCCAGCATGAAACCCAGGGAATGTCAGCAATTCAGTAAACGCAAA 182
QY 756 CAGAGCTGGAAGGTTCTAAAGAAATCTTAGTTCCTCAGAGGATGGCCCGCAGTATGGA 815

```

```

Db 181 CAGAGCTGGAAGGTTCTTAAAGAACTCTTAGTTCTCAGAGGCGATGGCCGCCACTAGGA 122
Qy 816 GAAAGTGGCCCTATCATCTGAGAGTCCCACTCAGACAGATGATAGGAGACCTTTG 875
Db 121 GAAAGTGGCCCTATCATCTGAGAGTCCCACTCAGACAGATGATAGGAGACCTTTG 62
Qy 876 TCTGATCTGCTATATCAGCATTTGTTCAAGCTCACCCTAGCAACGGAGAGAGTGAAC 935
Db 61 TCTGATCTGCTATATCAGCATTTGTTCAAGCTCACCCTAGCAACGGAGAGAGTGAAC 2
Qy 936 C 936
Db 1 C 1

RESULT 4
AV708169 694 bp mRNA linear EST 09-OCT-2000
DEFINITION AV708169 ADC Homo sapiens cDNA clone ADCALE11 5', mRNA sequence.
ACCESSION AV708169
VERSION AV708169.1 GI:10725434
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 694)
Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao,
H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu,
G., Hu, R., Chen, J., Chen, Z. and Han, Z.
GenBank accession number: U000000000
This clone is available at CHGC in Shanghai.
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
201203, P. R. China
Tel: 86-21-50801919 (ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
This clone is available at CHGC in Shanghai.
Location/Qualifiers
1. 694
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADCALE11"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT 229 a 177 c 124 g 162 t 2 others
ORIGIN
Query Match: 16.3%; Score 650.8; DB 10; Length 694;
Best Local Similarity 97.8%; Pred. No. 3.5e-156;
Matches 680; Conservative 0; Mismatches 13; Indels 2; Gaps 2;
Qy 1477 GTCATACATGGAAGTATATACATGATGCTGAGCTGCAACAAATTTGGCTCAGT 1536
Db 1 GTCATACATGGAAGTATATACATGATGCTGAGCTGCAACAAATTTGGCTCAGT 60
Qy 1537 CATCTTCTGCTCCAGCCAGTCTATCCCGGAGCAGTGGGCGATGAATCCCTACATGGCAAC 1596
Db 61 CATCTTCTGCTCCAGCCAGTCTATCCCGGAGCAGTGGGCGATGAATCCCTACATGGCAAC 120
Qy 1597 CTAACATACAGTGAAGTCAAAATTTTATACCCATTAGGAAATAGAGCAAGTCACT 1656
Db 121 CTAACATACAGTGAAGTCAAAATTTTATACCCATTAGGAAATAGAGCAAGTCACT 180
Qy 1657 CTGCTAATCTCAGCTTCTCTCAACAGTCACTGCAATATACACAGCTCCTATTCAAC 1716

```

```

Db 181 CTGCTAATTTCTCAGCCTTCCGCTACAAAGTCACTGCAATTACACAGCTCCTATTCAAC 240
Qy 1717 ACCCTGTCAAAAGTATGCTGCTATTAAACCCAGAGCTACAGAGCTGCTTTTACGACCTACAC 1776
Db 241 AGCCTGTGAAAGATGCTGCTATTAAACCCAGAGCTACAGAGCTGCTTTTACGACCTACAC 300
Qy 1777 ACCCTTCTGGATACCAAGCAATTTCAAACTGTTCACACAGTCCCTTTCTCTGAGGAA 1836
Db 301 ACCCTTCTGGATACCAAGCAATTTCAAACTGTTCACACAGTCCCTTTCTCTGAGGAA 360
Qy 1837 CGCTTCAAAATGTGCTGATGCCACCTGTTCGAGCTCCAAACTATCAAGGACCAAC 1896
Db 361 CGCTTCAAAATGTGCTGATGCCACCTGTTCGAGCTCCAAACTATCAAGGACCAAC 420
Qy 1897 CACCACTTACCAAAACATCTGCTGACCAAAACCCATCTGTTCTCTCATACGAGTCAA 1956
Db 421 CACCACTTACCAAAACATCTGCTGACCAAAACCCATCTGTTCTCTCATACGAGTCAA 480
Qy 1957 TCAGTAAGCCTAGCAAGAGGATCAGCAAGCTTCGCCCAAGAGATGAGAGTGAAGA 2016
Db 481 TCAGTAAGCCTAGCAAGAGGATCAGCAAGCTTCGCCCAAGAGATGAGAGTGAAGA 540
Qy 2017 GTATGAAATCTGATAGTGGGATAAAGAAAGAAACAGATTTACAACTTCCACTATTA 2076
Db 541 GTATGAAATCTGATAGTGGGATAAAGAAAGAAACAGATTTACAACTTCCACTATTA 599
Qy 2077 CTCTTAGGAAAACAAAGAGATGAAGCGAGGAGGATCTCGTATTCAAGTTATTCTC 2136
Db 600 CTCTTAGGAAAACAAAGAGATGAAGCGAGGAGGATCTCGTATTCAAGTTATTCTC 659
Qy 2137 CTCAAGCATTTAAATTTCTTTATGGAGCAACATGTA 2171
Db 660 CTCAAGCA-TTAAATTTCTTTATGGAGCCACATGTA 693

RESULT 5
BM971435/c 657 bp mRNA linear EST 21-MAR-2002
LOCUS UI-CF-DUI-abd-1-16-0-UI.s2 UI-CF-DUI Homo sapiens cDNA clone
DEFINITION UI-CF-DUI-abd-1-16-0-UI 3', mRNA sequence.
ACCESSION BM971435
VERSION BM971435.1 GI:19589022
KEYWORDS EST
SOURCE human
ORGANISM Homo sapiens
REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
AUTHORS Mammalia; Euthera; Primates; Catarrhini; Homnidae; Homo.
1 (bases 1 to 657)
Bonaldo, M.F., Lennon, G. and Soares, M.B.
Normalization and subtraction: two approaches to facilitate gene
discovery
Genome Res. 6 (9), 791-806 (1996)
JOURNAL 97044477
MEDLINE Contact: McCray, PB
COMMENT McCray Lab
University of Iowa
2024 University of Iowa Med Labs, Iowa City, IA 52242, USA
Tel: 319 356 4866
Fax: 319 356 7171
Email: paul-mccray@uiowa.edu
Tissue Procurement: Dr. M. J. Welsh, University of Iowa
CDNA Library preparation: Dr. M. Bento Soares, University of Iowa
CDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: Researchers may obtain clones from Research
Genetics (www.resgen.com).
The following repetitive elements were found in this cDNA
sequence: 1-30, >AT-rich#low_complexity
Seq primer: M13 FORWARD
POLYA=Yes.
Location/Qualifiers
1. 657
/organism="Homo sapiens"

```

/db_xref="taxon:9606"
/clone="UI-CF-DUI-abd-1-16-0-UI"
/clone_lib="UI-CF-DUI"
/tissue_type="Primary Lung Epithelial Cells"
/dev_stage="adult"
/lab_host="DH10B (Life Technologies) (T1 phage resistant)"
/note="Organ: Lung; Vector: p7T3-Pac (Pharmacia) with a
modified polylinker; Site_1: Ecor I; Site_2: Not I;
UI-CF-DUI is a normalized cDNA library containing the
following tissue(s): Primary Lung Epithelial Cells. The
library was constructed according to Bonaldo, Lennon and
Soares, Genome Research, 6:791-806, 1996. First strand
cDNA synthesis was primed with an oligo-dT primer
containing a Not I site. Double stranded cDNA was ligated
to an Ecor I adaptor, digested with Not I, and cloned
directionally into p7T3-Pac vector. The oligonucleotide
used to prime the synthesis of first-strand cDNA contains
a library tag sequence that is located between the Not I
site and the (dT)18 tail. The sequence tag for this
library is GGCTGTAGGC.
TAG_LIB=UI-CF-DUI
TAG_TISSUE=Lung Epithelial Cells Tissue nos 359-368
TAG_SEQ=GGCTGTAGGC"

BASE COUNT 155 a 114 c 150 g 237 t 1 others
ORIGIN

Query Match 14.8%; Score 590.6; DB 14; Length 657;
Best Local Similarity 96.08; Pred. No. 1-le-140;
Matches 605; Conservative 0; Mismatches 25; Indels 0; Gaps 0;

Qy 1651 GTCACCTGCTAAATCTCAGCCTTCTGTACACAGTCACCTGCAATTACACAGCTCCTA 1710
Db 641 GTCACCTGCTAAATCTCAGCCTTCCGNTACACAGTCACCTGCAATTACACAGCTCCTA 582
Qy 1711 TTCACAGCCTGTGAAGATGCGGTATTAAACACAGCTACAGACTGCTTTAGCAC 1770
Db 591 TTCACAGCCTGTGAAGATGCGGTATTAAACACAGCTACAGACTGCTTTAGCAC 522
Qy 1771 CTACACACCCCTTTGGATACACAGCAATTCAACTGTTCAACCCAGTCTTTTCCTG 1830
Db 521 CTACACACCCCTTTGGATACACAGCAATTCAACTGTTCAACCCAGTCTTTTCCTG 462
Qy 1831 AGGAACCGCTTCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAAACATCAAG 1890
Db 461 AGGAACCGCTTCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAAACATCAAG 402
Qy 1891 GACCACAGCACCCTACCCAAACATCTGCTGCACCAAAACCCATCTCTCTCCATACG 1950
Db 401 GACCACAGCACCCTACCCAAACATCTGCTGCACCAAAACCCATCTCTCTCCATACG 342
Qy 1951 AGTCAATCAGTAAGCCTAGCAAGAGGATCAGCAAGCTTCCCAAGGAAGATGAGAGTG 2010
Db 341 AGTCAATCAGTAAGCCTAGCAAGAGGATCAGCAAGCTTCCCAAGGAAGATGAGAGTG 282
Qy 2011 AAAAGAGTATGAAATTTGATAGTGGGATAAAGAAAGAAAGAAAGAAAGAAAGAAAGTTCAC 2070
Db 281 AAAAGAGTATGAAATTTGATAGTGGGATAAAGAAAGAAAGAAAGAAAGAAAGTTCAC 222
Qy 2071 CTATTACTGTTAGAAAACAAAGAAAGATGAAGACGAGGAAATCTCGTATTCAAGATT 2130
Db 221 CTATTACTGTTAGAAAACAAAGAAAGATGAAGACGAGGAAATCTCGTATTCAAGATT 162
Qy 2131 ATTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTC 2190
Db 161 ATTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAATGTACTCAAAATCTC 102
Qy 2191 ATCAGACGCTTACATCGTAAAAAACAATTTAGAAATGAATGATCGGGTTGGATTAT 2250
Db 101 ATCAGACGCTTACATCGTAAAAAACAATTTAGAAATGAATGATCGGGTTAAACCTT 42
Qy 2251 CTCAGATGCCAGGATCAATGAGAAGA 2280
Db 41 TTAATAATGTCCATTTTATATACTTAAAAAA 12

RESULT 6
BG331628
LOCUS
DEFINITION 602432628F1 NIH_MGC_18 Homo sapiens CDNA clone IMAGE:4550430 5',
mRNA sequence.
ACCESSION BG331628
VERSION BG331628.1 GI:13138066
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 1203)
AUTHORS NIH-MGC http://mgi.nci.nih.gov/
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: DCTD/BTP/Gazdar
CDNA Library Preparation: Ling Hong/Rubin Laboratory
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L10M243 row: 9 column: 07
High quality sequence stop: 724.
FEATURES
Location/Qualifiers
1..1203
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:4550430"
/clone_lib="NIH_MGC_18"
/tissue_type="large cell carcinoma"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
EcoRI; CDNA made by oligo-dT priming. Directionally cloned
into EcoRI/XhoI sites using the following 5' adaptor:
GGCAGAG(G). Library constructed by Ling Hong in the
laboratory of Gerald M. Rubin (University of California,
Berkeley) using ZAP-cDNA synthesis kit (Stratagene) and
Superscript II RT (Life Technologies). Note: this is a
NIH_MGC Library."
BASE COUNT 330 a 272 c 281 g 320 t
ORIGIN
Query Match 14.1%; Score 560; DB 12; Length 1203;
Best Local Similarity 98.3%; Pred. No. 1e-132;
Matches 566; Conservative 0; Mismatches 10; Indels 0; Gaps 0;
Qy 3 CTTTGGGTTGTTGGACGGACTCTGGCGCTCAGCGTCCGCCCTCAGGCCCGCGGCGGC 62
Db 90 CTTTGGGTTGTTGGACGGACTCTGGCGCTCAGCGTCCGCCCTCAGGCCCGCGGCGGC 149
Qy 63 TGTCCAGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTATTTATTTATTTAAAC 122
Db 150 TGTCCAGAGCTCTGCTCTCCCTCCAGAGTAAATTTATTTATTTATTTATTTAAAC 209
Qy 123 AGTCTTGGGACTTCTCTTTGAAGGATCAATTTTCACTTTTTCCTCAGAAGAAAGCTCTGGATC 182
Db 210 AGTCTTGGGACTTCTCTTTGAAGGATCAATTTTCACTTTTTCCTCAGAAGAAAGCTCTGGATC 269
Qy 183 TATCAATAAAGAAGTCTCTCGTGTGGCTACATATATATATATATATATATATATATATATAT 242
Db 270 TATCAATAAAGAAGTCTCTCGTGTGGCTACATATATATATATATATATATATATATATATAT 329
Qy 243 GAAAGCCAGAGGATATAGACAAATGAGGCTTAAGACCTTTCTGCCAGTAATCTACT 302
Db 330 GAAAGCCAGAGGATATAGACAAATGAGGCTTAAGACCTTTCTGCCAGTAATCTACT 389
Qy 303 GTCAGTAGCCGCAATGTTTACAAGAAATTCGGGAATTCCTTAGGAATTTATCTAAACCA 362

9
10
11
12

[illegible]

EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 597)

REFERENCE AUTHORS

Nagal, M. A.,
Goldman, G. H.,
Brunstein, A.,
M. J., Soares,
Simpson A. T.

Simpson, A.O.: Shotgun sequencing of the human transcriptome with ORF expressed sequence tags
Proc. Natl. Acad. Sci. U.S.A. 97 (7), 3491-3496 (2000)
20202663
Contact: Simpson A.J.G.
Laboratory of Cancer Genetics
Ludwig Institute for Cancer Research

Tel: +55-11-2704922
 Fax: +55-11-2707001
 Email: asimpsoneludwig.org.br

Email: asimpson@u.wyoming.edu
This sequence was derived from the FAPESP/LICR Human Cancer Genome

Project. This entry can be seen in the following URL
(<http://www.ludwig.org.br/scripts/gethtml2.pl?t1=MR0&t2=MR0-BT4508->

060701-003-h06&t3=2001-07-06&t4=1)

High quality sequence start: 10

High quality sequence stop: 596.

Location/Qualifiers
1 597

```
1. 1.557
/organism="Homo sapiens"
```

```

/db_xref="taxon:9606"

```

```
/clone_lib="BT4508"
/Dev stage="adult"
```

```

/dev_stayc= Adult
/note="Organ: breast; Vector: puc18; Site_1: SmaI; Site_2:

```

SmaI; A mini-library was made by cloning products derived

from ORESTES PCR (U.S. Letters Patent application No. 196 716 - Institute for Cancer Research) and pGL3- β -galactosidase.

into the pUC 18 vector. Reverse transcription of tissue

mRNA and cDNA amplification were performed under low

208 a	117 c	125 c	147 +
stringency conditions.			

200 a 11/ 122 14/ 1

ch	Score	DB	Length
13.88	548.2	14	597

Similarity 99.0%; Pred: NO. 8.3e-130;
83: Conservative 0: Mismatches 3: Indels 3: Gaps 3:

1. The first step is to identify the problem or question that needs to be answered. This involves understanding the context and the specific requirements of the task.

TTACATATATAGATGTTTTCATGA - AGAGGAGTGAAAAGCCAGAGGATATAGACAAAT 268

TACATATATAGATGTTTTCATGCACAGAGCGAGTCAAAAGCCACAGACCATATACACAAT 69

CC BY-NC-ND 4.0 International license.

AGGCCTAAGACCTTTCCTGCCAGTAACTATACTGTCAGTAGCCGCAATGTTACAAGA 328

CG-CTACACCTTCTCCACCTAACCTATCTCTCCACCTTCTACAC 138

120 110 100 90 80 70 60 50 40 30 20 10 0

ATTCTGGGAATCCCTTAGGAATTATCTAAACCATCTGATGCTAGGCTGAGCATAA 388

1	0
2	0
3	0
4	0
5	0
6	0
7	0
8	0
9	0
10	0
11	0
12	0
13	0
14	0
15	0
16	0
17	0
18	0
19	0
20	0
21	0
22	0
23	0
24	0
25	0
26	0
27	0
28	0
29	0
30	0
31	0
32	0
33	0
34	0
35	0
36	0
37	0
38	0
39	0
40	0
41	0
42	0
43	0
44	0
45	0
46	0
47	0
48	0
49	0
50	0
51	0
52	0
53	0
54	0
55	0
56	0
57	0
58	0
59	0
60	0
61	0
62	0
63	0
64	0
65	0
66	0
67	0
68	0
69	0
70	0
71	0
72	0
73	0
74	0
75	0
76	0
77	0
78	0
79	0
80	0
81	0
82	0
83	0
84	0
85	0
86	0
87	0
88	0
89	0
90	0
91	0
92	0
93	0
94	0
95	0
96	0
97	0
98	0
99	0
100	0

ATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCTAAGGCTGAGCAATA 100

ATGAGTAAATGTCAACCGAAGATCCTCGACAAGTCAGAAATCCACCCAAATTGGGAC 448

[illegible]

ATGAGTAAATGTCAACCGAAGATCCTCGACAAGTCCAGAAATCCACCCAAATTTGGGAC 248

ATCATAAAGCCTTGCAGGAAATTCGAACTCTGCTTCCATTGCAATGAAACAA 508

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

CAATTATAAAGCCTTGCAGGAAATTCGAAACTCTCTGCTTCCATTTGCAATGAAACAA 308

CTTCTCGGAGTACTTCAGAAAGTTAATCCACAAATGCTTCAAGACTTGC-AAGCTGCTG 567

1
2
3
4
5
6
7
8
9
10
11
12
13
14
15
16
17
18
19
20
21
22
23
24
25
26
27
28
29
30
31
32
33
34
35
36
37
38
39
40
41
42
43
44
45
46
47
48
49
50
51
52
53
54
55
56
57
58
59
60
61
62
63
64
65
66
67
68
69
70
71
72
73
74
75
76
77
78
79
80
81
82
83
84
85
86
87
88
89
90
91
92
93
94
95
96
97
98
99
100
101
102
103
104
105
106
107
108
109
110
111
112
113
114
115
116
117
118
119
120
121
122
123
124
125
126
127
128
129
130
131
132
133
134
135
136
137
138
139
140
141
142
143
144
145
146
147
148
149
150
151
152
153
154
155
156
157
158
159
160
161
162
163
164
165
166
167
168
169
170
171
172
173
174
175
176
177
178
179
180
181
182
183
184
185
186
187
188
189
190
191
192
193
194
195
196
197
198
199
200
201
202
203
204
205
206
207
208
209
210
211
212
213
214
215
216
217
218
219
220
221
222
223
224
225
226
227
228
229
230
231
232
233
234
235
236
237
238
239
240
241
242
243
244
245
246
247
248
249
250
251
252
253
254
255
256
257
258
259
260
261
262
263
264
265
266
267
268
269
270
271
272
273
274
275
276
277
278
279
280
281
282
283
284
285
286
287
288
289
290
291
292
293
294
295
296
297
298
299
300
301
302
303
304
305
306
307
308
309
310
311
312
313
314
315
316
317
318
319
320
321
322
323
324
325
326
327
328
329
330
331
332
333
334
335
336
337
338
339
340
341
342
343
344
345
346
347
348
349
350
351
352
353
354
355
356
357
358
359
360
361
362
363
364
365
366
367
368
369
370
371
372
373
374
375
376
377
378
379
380
381
382
383
384
385
386
387
388
389
390
391
392
393
394
395
396
397
398
399
400
401
402
403
404
405
406
407
408
409
410
411
412
413
414
415
416
417
418
419
420
421
422
423
424
425
426
427
428
429
430
431
432
433
434
435
436
437
438
439
440
441
442
443
444
445
446
447
448
449
450
451
452
453
454
455
456
457
458
459
460
461
462
463
464
465
466
467
468
469
470
471
472
473
474
475
476
477
478
479
480
481
482
483
484
485
486
487
488
489
490
491
492
493
494
495
496
497
498
499
500
501
502
503
504
505
506
507
508
509
510
511
512
513
514
515
516
517
518
519
520
521
522
523
524
525
526
527
528
529
530
531
532
533
534
535
536
537
538
539
540
541
542
543
544
545
546
547
548
549
550
551
552
553
554
555
556
557
558
559
560
561
562
563
564
565
566
567
568
569
570
571
572
573
574
575
576
577
578
579
580
581
582
583
584
585
586
587
588
589
590
591
592
593
594
595
596
597
598
599
600
601
602
603
604
605
606
607
608
609
610
611
612
613
614
615
616
617
618
619
620
621
622
623
624
625
626
627
628
629
630
631
632
633
634
635
636
637
638
639
640
641
642
643
644
645
646
647
648
649
650
651
652
653
654
655
656
657
658
659
660
661
662
663
664
665
666
667
668
669
670
671
672
673
674
675
676
677
678
679
680
681
682
683
684
685
686
687
688
689
690
691
692
693
694
695
696
697
698
699
700
701
702
703
704
705
706
707
708
709
710
711
712
713
714
715
716
717
718
719
720
721
722
723
724
725
726
727
728
729
730
731
732
733
734
735
736
737
738
739
740
741
742
743
744
745
746
747
748
749
750
751
752
753
754
755
756
757
758
759
760
761
762
763
764
765
766
767
768
769
770
771
772
773
774
775
776
777
778
779
780
781
782
783
784
785
786
787
788
789
790
791
792
793
794
795
796
797
798
799
800
801
802
803
804
805
806
807
808
809
810
811
812
813
814
815
816
817
818
819
820
821
822
823
824
825
826
827
828
829
830
831
832
833
834
835
836
837
838
839
840
84

```

Db 309 TTCTTCTCGGAGTACTTCAGAGTAAATCCACAATGCTTCAAGACTTGCAGAAAGCTGCTG 368
QY 568 GATTGTGATGAGGATATGTTATACAAAGCTTTCAGAAAACCTTAACAACAGAGATATAGAAG 627
Db 369 GATTGTGATGAGGATATGTTATACAAAGCTTTCAGAAAACCTTAACAACAGAGATATAGAAG 428
QY 628 CAGCAATGTAATTCATTAGTAAATAGTATACCAAGATCTTCGACGAGAGACAGATGCTG 687
Db 429 CAGCAATGTAATTCATTAGTAAATAGTATACCAAGATCTTCGACGAGAGACAGATGCTG 488
QY 688 CAGCAGTCCGACAGCTTATTAATCCAGCATGAACAGGAGATGTCAGCAATCACTTA 747
Db 489 CAGCAGTCCGACAGCTTATTAATCCAGCATGAACAGGAGATGTCAGCAATCACTTA 548
QY 748 ACCGCAAAACAGAGCTGGAAGGTTCTAAGAAATCCTTAGTTCCTCAGAG 796
Db 549 ACCGCAAAACAGAGCTGGAAGGTTCTAAGAAATCCTTAGTTCCTCAGAG 597

RESULT 10
AW502306 550 bp mRNA linear EST 01-MAR-2000
LOCUS UI-HF-BR0p-ajs-c-10-0-UI.r1 NIH_MGC_52 Homo sapiens cDNA clone
DEFINITION IMAGE:3075474 5', mRNA sequence.
ACCESSION AW502306
VERSION AW502306.1 GI:7116847
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 550)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: Louis M. Staudt, M.D., Ph.D.
cDNA Library Preparation: M.B. Soares Lab
cDNA Library Arrayed by: M.B. Soares Lab
DNA Sequencing by: M.B. Soares Lab
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
www-bio.llnl.gov/bbrp/image/image.html
Seq primer: M3 Forward
Location/Qualifiers
1..550
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3075474"
/clone_lib="NIH_MGC_52"
/tissue_type="lymph"
/cell_type="germinal center B cells"
/lab_host="MGC85"
/notes="vector: p773-Pac; Site_1: NotI; Site_2: Eco RI;
Constructed from size fractionated cytoplasmic mRNA
(7.4-9.5kb). Directionally cloned. Cells provided by
Louis M. Staudt, Ph.D. Library preparation by Maria de
Fatima Bonaldo, Ph.D. and M. Bento Soares, Ph.D."
BASE COUNT 187 a 120 c 117 g 126 t
ORIGIN
Query Match 13.6%; Score 543.2; DB 10; Length 550;
Best Local Similarity 99.5%; Pred No 1.6e-128;
Matches 545; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

QY 315 CAAATGTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCT 374
Db 3 CGAGGGTTACAGAAATTCGGGAATCCCTTAGGAATTTATCTAAACCATCTGATGCTGCT 62
QY 375 AAGGCTGAGCATACATGATGATTAATGTCAACCAAGATCTTCGACAGAGTCAAGAAATCCA 434

```

```

Db 63 AAGGCTGAGCATACATGATGATTAATGTCAACCAAGATCTTCGACAGAGTCAAGAAATCCA 122
QY 435 CCCAAATTTGGGACCGCATCATAAAGCTTTCGAGAAATTCGAAACTCTCTGCTTCCATT 494
Db 123 CCCAAATTTGGGACCGCATCATAAAGCTTTCGAGAAATTCGAAACTCTCTGCTTCCATT 182
QY 495 GCAGAAATGAACAAATTCCTTCGGAGTACTTTCAGAAATTAATCCCAAAATGCTTCAAGAC 554
Db 183 GCAATGAACAAATTCCTTCGGAGTACTTTCAGAAATTAATCCCAAAATGCTTCAAGAC 242
QY 555 TTGCAAGCTGCTGGATTTGATGAGGATATGTTATACAAAGCTTTCAGAAACTTCAACAC 614
Db 243 TTGCAAGCTGCTGGATTTGATGAGGATATGTTATACAAAGCTTTCAGAAACTTCAACAC 302
QY 615 AGAAGTATAGAACGACCAATTTGAATTCATTAGTAAATGAGTTACCAAGATCCTCGACGA 674
Db 303 AGAAGTATAGAACGACCAATTTGAATTCATTAGTAAATGAGTTACCAAGATCCTCGACGA 362
QY 675 GAGCAGATGGCTGCAGAGCTGCCAGACCTTATTAATGCCAGATGAACACGAGGATGTG 734
Db 363 GAGCAGATGGCTGCAGAGCTGCCAGACCTTATTAATGCCAGATGAACACGAGGATGTG 422
QY 735 CAGCAATCAGTTAACCCGCAACAGAGCTGGAAGGTTCTAAAGAAATCCTTAGTTCCTCAG 794
Db 423 CAGCAATCAGTTAACCCGCAACAGAGCTGGAAGGTTCTAAAGAAATCCTTAGTTCCTCAG 482
QY 795 AGGCATGGCCGACCTAGGAGAAAGTGTGGCTTATCTCTGAGAGTCCCAACTCACAG 854
Db 483 AGGCATGGCCGACCTAGGAGAAAGTGTGGCTTATCTCTGAGAGTCCCAACTCACAG 542
QY 855 ACAGATGT 862
Db 543 ACAGATGT 550

RESULT 11
BE568367 855 bp mRNA linear EST 15-AUG-2000
LOCUS 601341939F1 NIH_MGC_53 Homo sapiens cDNA clone IMAGE:3684273 5',
DEFINITION mRNA sequence.
ACCESSION BE568367
VERSION BE568367.1 GI:9812087
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 855)
NIH-MGC http://mgi.nci.nih.gov/.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabbs-remail.nih.gov
Tissue Procurement: ATCC
cDNA Library Preparation: CLONTECH Laboratories, Inc.
DNA Sequencing by: The I.M.A.G.E. Consortium (LLNL)
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at: image.llnl.gov
Plate: L1CM370 row: m column: 10
High quality sequence stop: 593.
Location/Qualifiers
1..855
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3684273"
/clone_lib="NIH_MGC_53"
/tissue_type="carcinoma, cell line"
/lab_host="DH10B (T1 phage-resistant)"
/notes="Organ: bladder; Vector: pDNR-LIB (Clontech);
Site_1: SfiI (ggccctcgcc); Site_2: SfiI (ggccattatggcc
); Double-stranded cDNA was prepared from cell line RNA."
FEATURES
source

```

RESULT	12
LOCUS	BB557112
DEFINITION	673 bp mRNA linear EST 26-OCT-2001 BB557112 RIKEN full-length enriched, 2 days pregnant adult female ovary Mus musculus cDNA clone E330026M13.3', mRNA sequence.
ACCESSION	BB557112
VERSION	BB557112.2 GI:16448760
KEYWORDS	EST.
SOURCE	house mouse.
ORGANISM	Mus musculus
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus... (bases 1 to 673)

GAGAGAGAGATTCTCGAGTTAATTAAATATCCCCCCCCCCC 3'. cDNA primer adapter of sequence [5

BASE COUNT 211 a 157 c 152 g 153 t

ORIGIN

Query Match 13.3%; Score 530.6; DB 10; Length 673;

Best Local Similarity 86.8%; Pred. No. 3e-125;

Matches 584; Conservative 0; Mismatches 89; Indels 0; Gaps 0;

QY 1720 CTGTGAAAGATGCGGTGATTAAACACGAGCTACAGAGCTGCTTTAGCACCTACACAC 1779

Db 1 CCGTGAAGACATGGCGTCTGAACACGAGCTGAGAGCTGTTTAGCCCAACCCATC 60

QY 1780 CTCTTGGATACACACCAATTCACACTGTCACCCAGCTCTTTCTGAGGGAACCG 1839

Db 61 CTCTTGAATGCCACACCAAGTTCAGACTGTTGAGCTACCCCTTTCTGAGGGTACAG 120

QY 1840 CTTCAATGTCAGTGTGATGCCACCTGCTGAAGCTCCAACTATCAAGGACCCAC 1899

Db 121 CTTCAAGTGTGCTGTGATCCACCTGCTGAGCTCCAGCTATCAAGGTCACCCAC 180

QY 1900 CACCTTACCCAAAATCTGCTGACCAACCAACCACTGTTCTCCATAGCAGTCAATCA 1959

Db 181 CGCTTATCCAAAATCTGCTGACCAACCAACCACTGTTCTCCATAGCAGTCAATCA 240

QY 1960 GTAGCCCTAGCAAGAGGATCAGCAAGCTGCCCAGGAGAGTGAAGTGAAGAGTT 2019

Db 241 GTAGCCCTGCAAGAGTGAAGAGGATGCTTACCCAGGAGAGTGAAGTGAAGAGTT 300

QY 2020 ATGAAATGTTGATGAGTGGGATAAAGAAAGAAAGAGATTCACACTTACCTATTACTG 2079

Db 301 CGGACAGTGGTGACTCTGGGGATAAAGAAAGAAAGAGATTCACACTTACCTATTACTG 360

QY 2080 TTAGGAAAACCAAGAGATGAGAGGAGGATCTCGTATTCAAGTATTCTCTCTC 2139

Db 361 TTGCGAAAACCAAGAGATGAGAGGAGGATCTCGGATTCAGAGTTACTTCCCTAC 420

QY 2140 AGCATTTAAATCTTTTATGAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGC 2199

Db 421 AGGCTTTAAATCTTTTATGAGCAACATGTAGAAAATGTACTCAAAATCTCATCAGCAGC 480

QY 2200 GTCTACATCGTAAAAACAATAGAGATGAATGATGATGATGATGATGATGATGATGATG 2259

Db 481 GTCTACATCGTAAAAACAATAGAGATGAATGATGATGATGATGATGATGATGATGATG 540

QY 2260 CCGAGATCAATGAGAGAGTCTTTCCAAAAGAGATCTAAATATACATCGCTCTTAAA 2319

Db 541 CCGAGATCAATGAGAGAGTCTTTCCAAAAGAGATCTAAATATACATCGCTCTTAAA 600

QY 2320 GGGCTAAAATGGACAGTCTATGTTGTGAAGATAAAGACACTAGGATAGGACATTTG 2379

Db 601 GGGCTAAAATGGACAGTCTATGTTGTGAAGATAAAGACACTAGGATAGGACATTTG 660

QY 2380 GTGAAGTCTGTCT 2392

Db 661 GTGAAGTCTGTCT 673

RESULT 13

BI15504

LOCUS 648 bp mRNA linear EST 05-JUL-2001

DEFINITION 602903902F1 NCI_CGAP_Mam3 Mus musculus CDNA clone IMAGE:5033425 5', mRNA sequence.

ACCESSION BI15504

VERSION BI15504

KEYWORDS EST.

SOURCE house mouse.

ORGANISM Mus musculus

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. (bases 1 to 648)

AUTHORS NIH-MGC

TITLE NIH-MGC Institutes of Health, Mammalian Gene Collection (MGC)

JOURNAL Unpublished (1999)

COMMENT

Contact: Robert Strausberg, Ph.D.
Email: crabs-r@mail.nih.gov
Tissue Procurement: Lohar Hennighausen Ph.D., Chu-Xia Deng Ph.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLAM11093 row: d column: 02
High quality sequence stop: 648.

FEATURES

Location/Qualifiers

1. 648

/organism="Mus musculus"

/strain="129,C57BL/6J,FVB/N"

/db_xref="taxon:10090"

/clone="IMAGE:5033425"

/clone_lib="NCI_CGAP_Mam3"

/tissue_type="tumor, gross tissue"

/dev_stage="10 months"

/lab_host="DH10B"

/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI; Site_2: NotI; Cloned unidirectionally. Primer: Oligo dt. Library constructed by Life Technologies. Investigators providing samples: Lohar Hennighausen/Chu-Xia Deng, NIH Reference for transgenic model: Xu et al., Nature Genetics 22, 37-43 (1999)."

BASE COUNT 184 a 132 c 167 g 165 t

ORIGIN

Query Match 13.0%; Score 518.4; DB 13; Length 648;

Best Local Similarity 90.2%; Pred. No. 4.1e-122;

Matches 577; Conservative 0; Mismatches 61; Indels 2; Gaps 2;

QY 2377 TTGGTGAAGTCTCTAGCAAGAAAAGTAGATACTAAGCTTTGTATGCAACAAAACATC 2436

Db 11 TTGGTGAAGTCTCTAGCAAGAAAAGTAGATACTAAGCTTTGTATGCAACAAAACATC 59

QY 2437 TTCGAAAGAAAGATGTTCTTCTTCGAAATCAAGTCCCTCATGTTAAGGCTGAGAGAGATA 2496

Db 70 TTCGAAAGAAAGAGC-TCTGCTCCGAAATCAGGTGGCTCATGTGAAGCGGAGAGGATA 128

QY 2497 TCCTGGCTGAAGCTGACAATGAATGGTAGTTCGTCTATATTATTTCATTCACAGATAAG 2556

Db 129 TCCTAGCAGAGCGGACAAATGAGTGGTGTGCGCTGTACTACTCTTCTCAGGACAAG 188

QY 2557 ACAATTTATCTTTGTAATGGACTACATTCCTGGGGTGTATGATGAGCTTATTAATTA 2616

Db 189 ACAATTTGTAATGGACTACATTCCTGGGGTGTATGATGAGCTTATTAATTA 248

QY 2617 GAATGGGCATCTTCCAGAAAGCTGCGACGATTTCTACATAGCAGAACTTACCTGTGCAG 2676

Db 249 GAATGGGCATCTTCCAGAAAGCTGCGACGATTTCTACATAGCAGAACTTACCTGTGCAG 308

QY 2677 TTGAAAGTGTTCATATAAATGGGTTTATTTCATAGAGATATTAAACCTGATATTTTGA 2736

Db 309 TTGAAAGTGTTCATATAAATGGGTTTATTTCATAGAGATATTAAACCTGATATTTTGA 368

QY 2737 TTGATCGTGAAGTTCATATTAATTTGACTGACTTTGGCTCTGCACCTGGCTTCAGATGGA 2796

Db 369 TTGACCGTGAAGTTCATATTAATTTGACTGACTTTGGCTCTGCACCTGGCTTCAGATGGA 428

QY 2797 CACAGGATTTAAAGTACTTATCAGAGTGGTGACCACTCCACGCAAGATAGCATGATTTC 2856

Db 429 CACAGGATTTAAAGTACTTATCAGAGTGGTGACCACTCCACGCAAGATAGCATGATTTC 488

QY 2857 GTATGTAATGGGGATCCCTCAAGCTGTCAGTGGGAGAGAGTGAAGCCATTAGAGC 2916

Db 489 GTAAAGCAATGGGAGATCCCTCAATTTCTGGTGTGGGAGAGAGTGAAGCCATTAGAGC 548

QY 2917 GGAGAGCTGCAGCCGAGCAGCAGCAGCTGCTAGCAGATTTCTTTGTTGGGACTCCCAATT 2976

Db 549 GGAGAGCTGCAGCCGAGCAGCAGCAGCTGCTAGCAGATTTCTTTGTTGGGACTCCCAATT 608

```

QY 2977 ATATTGCACCTGAAGTGTGCTACGAACAGGATACACACA 3016
|||||
Db 609 ATATTGCACCTGAAGTGTGCTACGAACAGGATACACACA 648

RESULT 14
BG088986 658 bp mRNA linear EST 26-JAN-2001
LOCUS ut60q11.v1 Soares mouse NMGB.bcell Mus musculus cDNA clone
DEFINITION IMAGE:3332540 5' similar to TR:095835 095835 LARGE TUMOR SUPPRESSOR
1.. mRNA sequence.
ACCESSION BG088986
VERSION BG088986.1 GI:12571583
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 658)
NCI-CCAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-re@mail.nih.gov
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
MGI:1076704
Seq primer: -40RP from Gibco
High quality sequence stop: 473.
FEATURES
Location/Qualifiers
1..658
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3332540"
/lab_host="Soares mouse NMGB.bcell"
/note="Organ: germinal B-cell; Vector: p773D-Pac
(Pharmacia) with a modified polylinker; Site.1: Not.I;
Site.2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer 15'
TGTTACCAATCTGAAGTGAGGAGCGGCCCGCTGTTTTTTTTTTTTTTTTTTTT
T 3'; double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified p773 vector.
Library is normalized; constructed by Bento Soares and
M.Fatima Bonaldo."
BASE COUNT 180 a 206 c 136 g 135 t 1 others
ORIGIN
Query Match 12.8%; Score 511; DB 12; Length 658;
Best Local Similarity 86.7%; Pred. No. 3.3e-120;
Matches 562; Conservative 0; Mismatches 86; Indels 0; Gaps 0;

QY 578 GGATATGTTATACAGCTCTTCAGAAACTAACACAGAGATATAGAGCAGCAATTGA 637
|||||
Db 1 GGACATGTTATTCAAGCTCTTCAGAAACCAATACAGAACATAGAGCGCTGTTGA 60

QY 638 ATTCAATTAGTAATAGTTACCAAGATCTTCGACAGACAGATGGCTGCAGCAGCTGC 697
|||||
Db 61 ATTCAATTAGTAATAGTTACCAAGATCTTCGACAGACAGATGGCTGCAGCAGCTGC 120

QY 698 CAGACCTATTATCCAGCATCAACACAGGGAATGTGCAGCAATCAGTTAAACCGCAACA 757
|||||
Db 121 CAGGCCTATTATAGCCACCATCAACACAGGAAATGTCAACATTTCAATTAACCGAAACA 180

QY 758 GAGCTGGAAGGTTCTAAGAAATCTTAGTTCTCTCAGAGGATGGCCGCCACTAGGAGA 817
|||||
Db 181 AAGCTGGAAAGGTTCTAAGAGTCTCTAGTTCTCTCAGACACAGCGCCATCTCTAGGAGA 240
|||||
QY 818 AAGTGTGCGCTATATCTTGAGAGTGCCCAACTCAGACAGATGATGAGGAACCTTTGTC 877
|||||

```

241 AAATGTGGTTTATCGTTCTGAAAGCCCAACTCACAGCGGATGTAGGAAGACCTCTGTC 300

QY 878 TGSATCTGTATATACAGCATTTCTCAAGCTCACCTACAGCGGACAGAGTGAACCC 937
|||||

Db 301 TGAATCCGGCATTTGCAGCATTTGCTCAAGCTCACCAAGCAATGGACAGAGTGAACCC 360
|||||

QY 938 CCCACCACCACCTCAAGTAGAGGAGTGTACTCTCTCCACCACCTCCCAAGAGGCGCAGACTCC 997
|||||

Db 361 CCCACCACCACCTCAAGTTAGGAGTGTACTCTCTCCACCACCTCCCGAGAGGCGCAGACTCC 420
|||||

QY 998 CCCTCAAGAGGTACAACTCCACCTCCCTTCATCGGAGCAACAACTCTCAACAAGCG 1057
|||||

Db 421 ACCTCCCGAGGACCACTCCCTCCCTCCCTCATGGAAACCAAGCTCTCAGACAAGCG 480
|||||

QY 1058 CTATTCCTGGAACATGGAATAGTATCTCCGAAATCTCTCTGTCTCCACCTGGGCGATG 1117
|||||

Db 481 CTACTCTGGAACTAGGATAGTATCTCCGAAATCTCTCTGTCTCCACCTGGGCGATG 540
|||||

QY 1118 GCAAGAGGCTATCTCTCCACCACCTCTCAACACTTCCCTCCATGAATCTCTCTTAATCAAG 1177
|||||

Db 541 GCAGGAGGCTACCTCCACCACCTCTTACCACCTCTCTCCATGAATCTCTCCCTAGCCAGCG 600
|||||

QY 1178 ACAGAGAGCATAGTCTGTCTGTGGCAGACAAACCAATCATCAT 1225
|||||

Db 601 TCAGAGGCGCATAGTCTGTCTGTGGTGTAGTACTACCTCCACCAT 648
|||||

RESULT 15
AZ340286/c 665 bp DNA linear GSS 29-SEP-2000
LOCUS 1M0072J17F Mouse 10kb plasmid UUGC1M library Mus musculus genomic
DEFINITION clone UUGC1M0072J17 F, DNA sequence.

ACCESSION AZ340286
VERSION AZ340286.1 GI:10415231
KEYWORDS GSS.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 665)
Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
and Wright,D., Weiss,R.
Mouse whole genome scaffolding with paired end reads from 10kb
plasmid inserts
Unpublished (2000)
Contact: Robert B. Weiss
University of Utah Genome Center
University of Utah
Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
84112 USA
Tel: 801 585 5606
Fax: 801 585 7177
Email: ddunn@genetics.utah.edu
Insert Length: 10000 Std Error: 0.00
Plate: 0072 row: J column: 17
Seq primer: CGTTGTAACAGCAGCGCCAGT
Class: plasmid ends
High quality sequence stop: 665.
Location/Qualifiers
1..665
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGC1M0072J17"
/clone_lib="Mouse 10kb plasmid UUGC1M library"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, Tl-resistant, F-"
/note="Vector: PWD42hv; Purified genomic DNA from M.
musculus C57BL/6J (male) was obtained from the Jackson
Laboratory Mouse DNA Resource
(http://www.jax.org/resources/documents/dnares/). The DNA

was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of pWD42 (gi14732114|gb|AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor-mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT	162 a	132 c	204 g	167 t
ORIGIN				
Query Match	12.5%	Score 497.6;	DB 17;	Length 665;
Best Local Similarity	86.1%	Pred. No. 9;4e-117;		
Matches 575;	Conservative 0;	Mismatches 89;	Indels 4;	Gaps 2;
Qy 1160	GAATCTCTCTTAATCAAGACACAGAGGAGGATTAGTTCTGTTCTGTTGGCAGACACCA-A	1218		
Db		606		
Qy 1219	TCATCATGCAGAGTTCTAGCAAAATTTAACTTTCCATCAGGGAGACCTGGAATGCAGAAATG	1278		
Db		546		
Qy 1279	GTACTGGACAACTGATTTCATGATACACCAAAATGTTGCTCCTGCTGGCACCTGTGAATC	1338		
Db		489		
Qy 1339	GGCAGCACCACTCCATATCCTCTGACAGAGCTAATGGACAAAGCCCTTCTGCTTTAC	1398		
Db		429		
Qy 1399	AAACAGGGGGATCTGCTCTCTGCTCATATACAAATGGAAGTATCTCAGTCTATGA	1458		
Db		369		
Qy 1459	TGGTGCCAAACAGAAATAGTCATACATGGAACATATATAACATTAGTACCTGGACTGC	1518		
Db		309		
Qy 1519	AAACAAATGGCCCTCAGTCATCTCTGCTCCAGCCAGCTATCCCGAGAGTGGGCATG	1578		
Db		249		
Qy 1579	AAATCCCTTACATGGCAACCTAACATACCAGTGAGGTCAAAATCTTTTAAACCCATTAG	1638		
Db		189		
Qy 1639	GAATAGAGCAAGTCACTCTCTAATCTCAGCCCTTCTGCTACACAGTCACTGCAATTA	1698		
Db		129		
Qy 1699	CACCAGTCTCTTTCACAGAGCTGTGAAAAGTATGCGTGTATTAACACCAGAGCTACAGA	1758		
Db		69		
Qy 1759	CTGCTTTAGCACCTTACACACCTCTTCTGATACACAGCCCAATTCAACTGTTCAACCCA	1818		
Db		9		
Qy 1819	GTCTTTT 1826			
Db				
	8 CCCCTTTT 1			

GenCore version 5.1.3
Copyright (c) 1993 - 2003. Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 5334.88 Seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-3
Perfect score: 3213
Sequence: 1 gtgcaacattcaattaacg.....cccccaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

- 1: gb_ba.*
- 2: gb_htg.*
- 3: gb_in.*
- 4: gb_om.*
- 5: gb_ov.*
- 6: gb_pat.*
- 7: gb_ph.*
- 8: gb_pl.*
- 9: gb_pr.*
- 10: gb_ro.*
- 11: gb_sts.*
- 12: gb_sy.*
- 13: gb_un.*
- 14: gb_vi.*
- 15: em_ba.*
- 16: em_fun.*
- 17: em_hum.*
- 18: em_in.*
- 19: em_mu.*
- 20: em_om.*
- 21: em_or.*
- 22: em_ov.*
- 23: em_pat.*
- 24: em_ph.*
- 25: em_pl.*
- 26: em_ro.*
- 27: em_sts.*
- 28: em_un.*
- 29: em_vi.*
- 30: em_htg_hum.*
- 31: em_htg_inv.*
- 32: em_htg_other.*
- 33: em_htg_mus.*
- 34: em_htg_pln.*
- 35: em_htg_rod.*
- 36: em_htg_mam.*
- 37: em_htg_vrt.*
- 38: em_sy.*
- 39: em_htgo_hum.*
- 40: em_htgo_mus.*
- 41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	3213	100.0	3213	6	AR201456	AR201456 Sequence
2	3213	100.0	3213	10	AF104414	AF104414 Mus muscu
3	2483	77.3	3984	6	AR201455	AR201455 Sequence
4	2483	77.3	4307	9	AF104413	AF104413 Homo sapi
5	2472	76.9	4241	9	AF164041	AF164041 Homo sapi
6	1946.8	60.6	2442	6	E24613	E24613 warts Prote
7	1168.8	36.4	80146	9	AL583963	AL583963 Human DNA
8	1126.8	35.1	1374	6	E24614	E24614 warts Prote
9	1079.4	33.6	16490	2	AC102712	AC102712 Mus muscu
10	872.4	27.2	16490	2	AC102712	AC102712 Mus muscu
11	812.6	25.3	3155	6	AR201457	AR201457 Sequence
12	806.2	25.1	3460	10	AB023958	AB023958 Mus muscu
13	787.2	24.5	5486	6	E38226	E38226 Human tumor
14	787.2	24.5	5486	6	E38227	E38227 Human tumor
15	785.6	24.5	3412	9	AB028019	AB028019 Homo sapi
16	785.6	24.5	4098	9	AF207547	AF207547 Homo sapi
17	675.6	21.0	191604	2	AC126128	AC126128 Rattus no
18	609.8	19.0	181750	2	AC098623	AC098623 Rattus no
19	598.2	18.6	795	6	E24616	E24616 warts Prote
20	553	17.2	181750	2	AC098623	AC098623 Rattus no
21	521.2	16.2	5360	3	DR0WARTS	L39837 Drosophila
22	521.2	16.2	5720	3	DMU29608	U29608 Drosophila
23	521.2	16.2	5720	6	AR201454	AR201454 Sequence
24	468.4	14.6	582	6	E24615	E24615 warts Prote
25	353.4	11.0	183597	9	AL356285	AL356285 Human DNA
26	258.8	8.1	2422	8	AY128279	AY128279 Arabidops
27	254	7.9	236109	3	AE003775	AE003775 Drosophil
28	252.4	7.9	28133	2	AC012975	AC012975 Drosophil
29	252.4	7.9	163741	3	AC007821	AC007821 Drosophil
30	251.8	7.8	2202	8	MGPRTKINA	Z30329 M.crystalli
31	250	7.8	2355	8	SOPRTKINA	Z30330 S.oleracea
32	248.4	7.7	1935	6	AR139101	AR139101 Sequence
33	248.4	7.7	4725	9	BC028603	BC028603 Homo sapi
34	248.4	7.7	4983	6	AX056364	AX056364 Sequence
35	248.4	7.7	5181	9	AB023182	AB023182 Homo sapi
36	244.4	7.6	3018	6	A52140	A52140 Sequence 6
37	244.4	7.6	3018	6	AR084691	AR084691 Sequence
38	244.4	7.6	3018	9	HSPROKINX	Z35102 H.sapiens m
39	244.4	7.6	3593	9	BC012085	BC012085 Homo sapi
40	244	7.6	1745	3	AF275634	AF275634 Caenorhab
41	243	7.6	1738	8	AY096444	AY096444 Arabidops
42	243	7.6	2365	8	AY063969	AY063969 Arabidops
43	241.2	7.5	1322	3	E0C249683	AJ249683 Euplotes
44	241.2	7.5	1578	3	E0C249680	AJ249680 Euplotes
45	240.6	7.5	1804	3	E0C249682	AJ249682 Euplotes

ALIGNMENTS

RESULT 1
AR201456
LOCUS AR201456 3213 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 5 from patent US 6359193.
ACCESSION AR201456
VERSION AR201456.1 GI:20252344
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 3213)
AUTHORS Xu,T., Tao,W., Wang,W., Zhang,S. and Yu,W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 5 19-MAR-2002;
FEATURES Location/Qualifiers

```

source      1. .3213
BASE COUNT  946 a 764 c 696 g 807 t
ORIGIN

Query Match 100.0%; Score 3213; DB 6; Length 3213;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 3213; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GTGCAACATTCAATTAACCGAAACAAAGCTGGAAGGTTCTAAAGAGTCTCTAGTTCCT 60
DB 1 GTGCAACATTCAATTAACCGAAACAAAGCTGGAAGGTTCTAAAGAGTCTCTAGTTCCT 60
QY 61 CAGACACAGCCCATCTCTAGGAGAAATGTGTTTATCGTTCTGAAAGCCCAACTCA 120
DB 61 CAGACACAGCCCATCTCTAGGAGAAATGTGTTTATCGTTCTGAAAGCCCAACTCA 120
QY 121 CAGCGGATGTAGGAAGACCTCTCTGTGGATCCGGCATTCAGCATTTGCTCAAGCTCAC 180
DB 121 CAGCGGATGTAGGAAGACCTCTCTGTGGATCCGGCATTCAGCATTTGCTCAAGCTCAC 180
QY 181 CCAAGCAATGGACAGAGAGTGAACCCGCCACCCACCTCAAGTTAGGAGTGTACTCCT 240
DB 181 CCAAGCAATGGACAGAGAGTGAACCCGCCACCCACCTCAAGTTAGGAGTGTACTCCT 240
QY 241 CCACCACTCCGAGAGGCGAGACCCACCTCCCGAGGACCACTCCGCCCTCCGCCCTCA 300
DB 241 CCACCACTCCGAGAGGCGAGACCCACCTCCCGAGGACCACTCCGCCCTCCGCCCTCA 300
QY 301 TGGGAACCAAGCTCTCAGACAAAGCGTACTCTGGGACATGGAGTAACTCCCGA 360
DB 301 TGGGAACCAAGCTCTCAGACAAAGCGTACTCTGGGACATGGAGTAACTCCCGA 360
QY 361 ATCTCCCTGTTCCACCTGGGGCTGGCAGAGGGGTACCTCCACCACTCTTTACCACT 420
DB 361 ATCTCCCTGTTCCACCTGGGGCTGGCAGAGGGGTACCTCCACCACTCTTTACCACT 420
QY 421 TCTCCCATGAATCCCTTAGCAGGCTCAGAGGCCATTAAGTTCTGTTCAGTTGGTGA 480
DB 421 TCTCCCATGAATCCCTTAGCAGGCTCAGAGGCCATTAAGTTCTGTTCAGTTGGTGA 480
QY 481 CAACCCATCATGACAGAGTACTAGCAAAATTTAACTTTACACGAGGCGACCTGGAGTT 540
DB 481 CAACCCATCATGACAGAGTACTAGCAAAATTTAACTTTACACGAGGCGACCTGGAGTT 540
QY 541 CAGAAATGGTGTGCTAGTCTGATTTTATCGTGACCAAAATGTCCTGTTCTGTG 600
DB 541 CAGAAATGGTGTGCTAGTCTGATTTTATCGTGACCAAAATGTCCTGTTCTGTG 600
QY 601 ACTCGGAGCCACCACTCCATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 660
DB 601 ACTCGGAGCCACCACTCCATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 660
QY 661 TTACAAACAGGGCTCTGCTGCTCCACCATATTCCCAATGGAAGCTTCTCTAGTCTG 720
DB 661 TTACAAACAGGGCTCTGCTGCTCCACCATATTCCCAATGGAAGCTTCTCTAGTCTG 720
QY 721 ATGATGTGCCCACACAGGAAACAGTCAATGAGGCTTTATATATTAATGTCCTGGA 780
DB 721 ATGATGTGCCCACACAGGAAACAGTCAATGAGGCTTTATATATTAATGTCCTGGA 780
QY 781 CTGCAACAGCCTGGCCCCAGTGTCTTCTGCTCCGAGTCAATCCCAAGGGTGG 840
DB 781 CTGCAACAGCCTGGCCCCAGTGTCTTCTGCTCCGAGTCAATCCCAAGGGTGG 840
QY 841 CATGAATTCCTACATGGCAACCTAACATACAGTGAAGTCAAACTCTTTTAAACCCA 900
DB 841 CATGAATTCCTACATGGCAACCTAACATACAGTGAAGTCAAACTCTTTTAAACCCA 900
QY 901 TTAGGAAGTGAAGCAAGTCACTCTGCTAAATTTCTCAGCCTTCTGCGCACTACAGTGC 960
DB 901 TTAGGAAGTGAAGCAAGTCACTCTGCTAAATTTCTCAGCCTTCTGCGCACTACAGTGC 960

```

```

QY 961 ATCACACCGCTCTCTATTCAACACCGCTGAAAGAGCATGCGCGTCTCTGAAACAGAGCTG 1020
DB 961 ATCACACCGCTCTCTATTCAACACCGCTGAAAGAGCATGCGCGTCTCTGAAACAGAGCTG 1020
QY 1021 CAGACTGCTTTAGCCCAACCCCATCTCTTTGGATGCCACAGCCAGTTCAGACTGTTTCAG 1080
DB 1021 CAGACTGCTTTAGCCCAACCCCATCTCTTTGGATGCCACAGCCAGTTCAGACTGTTTCAG 1080
QY 1081 CCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCTCATCCACCTGTTGCTGAA 1140
DB 1081 CCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCTCATCCACCTGTTGCTGAA 1140
QY 1141 GCTCAAGCTATCAAGGTCCACACCGCTTTATCCAAAACATCTGTACACCAAAACCCA 1200
DB 1141 GCTCAAGCTATCAAGGTCCACACCGCTTTATCCAAAACATCTGTACACCAAAACCCA 1200
QY 1201 TCTGTCCCTCCATATGAGTCAAGTAAAGTAAAGCCCTGCAAGATGAACAGCTTAGCTTACC 1260
DB 1201 TCTGTCCCTCCATATGAGTCAAGTAAAGTAAAGCCCTGCAAGATGAACAGCTTAGCTTACC 1260
QY 1261 AAGGAAGATGATAGTGAAGAGTGGGACAGTGTGCTGCTGGGGATAAAGAAAGAAA 1320
DB 1261 AAGGAAGATGATAGTGAAGAGTGGGACAGTGTGCTGCTGGGGATAAAGAAAGAAA 1320
QY 1321 CAGATTACAACCTTCACCTATCATCTGTTTCGGAAGAACAGAAAGATGAAGAACAGAG 1380
DB 1321 CAGATTACAACCTTCACCTATCATCTGTTTCGGAAGAACAGAAAGATGAAGAACAGAG 1380
QY 1381 TCTCGGATTCAGAGTACTCCACACAGCTTTAAAGTCTTCTATGGAGCAGCAGTAGAG 1440
DB 1381 TCTCGGATTCAGAGTACTCCACACAGCTTTAAAGTCTTCTATGGAGCAGCAGTAGAG 1440
QY 1441 AAGCTCTGAACTCTCATCAGCAGCTGCTGATCGGAAGAACAGCAGCTAGAAAATGAAATG 1500
DB 1441 AAGCTCTGAACTCTCATCAGCAGCTGCTGATCGGAAGAACAGCAGCTAGAAAATGAAATG 1500
QY 1501 ATGCGGGTGGATTTATCTCAAGATGCCAGAGTCAAAATGAGAAAGATGCTTTGCCAGAAA 1560
DB 1501 ATGCGGGTGGATTTATCTCAAGATGCCAGAGTCAAAATGAGAAAGATGCTTTGCCAGAAA 1560
QY 1561 GAGTCTAACTATATTCCTCTTAAAGGGCTAAATGACAAAGTCTATGTTGTAAGATA 1620
DB 1561 GAGTCTAACTATATTCCTCTTAAAGGGCTAAATGACAAAGTCTATGTTGTAAGATA 1620
QY 1621 AAGACATTAGGAATAGGAGGCTTTGGTGAAGTCTGTCTAGCAAGAAAAGTCTGACTATAA 1680
DB 1621 AAGACATTAGGAATAGGAGGCTTTGGTGAAGTCTGTCTAGCAAGAAAAGTCTGACTATAA 1680
QY 1681 GCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGCTTCTGCTCCGAAATCAGGTGGCT 1740
DB 1681 GCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGCTTCTGCTCCGAAATCAGGTGGCT 1740
QY 1741 CATGTGAACCGGAGAGGGATATCTAGCAGAACCCACAAATGAGTGGGTGGTCCGCTG 1800
DB 1741 CATGTGAACCGGAGAGGGATATCTAGCAGAACCCACAAATGAGTGGGTGGTCCGCTG 1800
QY 1801 TACTACTCTTTCCAGGACAAAGCAACTTCTACTTTTGTGATGAGTACTTCTCTGGGGGG 1860
DB 1801 TACTACTCTTTCCAGGACAAAGCAACTTCTACTTTTGTGATGAGTACTTCTCTGGGGGG 1860
QY 1861 GATATGATGAGCCTATTAATAGAAAGGCTTCTTCTGAAATCTGGCAGCATTTCTAC 1920
DB 1861 GATATGATGAGCCTATTAATAGAAAGGCTTCTTCTGAAATCTGGCAGCATTTCTAC 1920
QY 1921 ATAGCAGAACTTACCTGTGAGTGAAGGTTGAAAGGTTTCAAAAATGGTCTTATTCATAGAGAT 1980
DB 1921 ATAGCAGAACTTACCTGTGAGTGAAGGTTGAAAGGTTTCAAAAATGGTCTTATTCATAGAGAT 1980
QY 1981 ATTAACCTGATAACTTTTATGATGAGCCTGATGGCCATATTAATTAATGACTGCTTGGC 2040
DB 1981 ATTAACCTGATAACTTTTATGATGAGCCTGATGGCCATATTAATTAATGACTGCTTGGC 2040
QY 2041 TTGTGCACTGGCTTCAGATGGACACATGACTCCAAAGTACTACCAAGTGGGGATCACCCA 2100

```

Db 3121 AATTCTTGGTAACTTAAAGTACTTAAAGAGAACCCCTGGTATCTCTTTTGTATATATAATAA 3180

QY 3181 ATAATTTTAAATCCCAAAAAA 3213

Db 3181 ATAATTTTAAATCCCAAAAAA 3213

RESULT 2

AF104414 3213 bp mRNA linear ROD 18-MAR-1999

LOCUS Mus musculus large tumor suppressor 1 (Lats1) mRNA, partial cds.

DEFINITION

ACCESSION AF104414

VERSION AF104414.1 GI:4324435

KEYWORDS

SOURCE Mus musculus.

ORGANISM Mus musculus

REFERENCE 1 (bases 1 to 3213)

AUTHORS Xu, T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.

TITLE Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase

JOURNAL Development 121 (4), 1053-1063 (1995)

MEDLINE 95262551

PUBMED 7743921

REFERENCE 2 (bases 1 to 3213)

AUTHORS Tao, W., Zhang, S., Turenchak, G.S., Stewart, R.A., St John, M.A., Chen, W. and Xu, T.

TITLE Human homologue of the Drosophila melanogaster lats tumour suppressor modulates CDC2 activity

JOURNAL Nat. Genet. 21 (2), 177-181 (1999)

MEDLINE 99140767

PUBMED 9988268

REFERENCE 3 (bases 1 to 3213)

AUTHORS St John, M.A., Tao, W., Fei, X., Fukumoto, R., Carcangiu, M.L., Brownstein, D.G., Parlow, A.F., McGrath, J. and Xu, T.

TITLE Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian tumours and pituitary dysfunction

JOURNAL Nat. Genet. 21 (2), 182-186 (1999)

MEDLINE 99140768

PUBMED 9988269

REFERENCE 4 (bases 1 to 3213)

AUTHORS Tao, W. and Xu, T.

TITLE Direct Submission

JOURNAL Submitted (04-NOV-1998) Genetics, HHMI, Yale University; 295 Congress Ave BCM 254D, New Haven, CT 06536, USA

FEATURES

source Location/Qualifiers

1..3213

/organism="Mus musculus"

/db_xref="taxon:10090"

/tissue_type="brain"

/dev_stage="newborn"

<1..3213

/gene="Lats1"

<1..2889

/gene="Lats1"

/note="putative serine threonine kinase; (MUM) Lats1"

/codon_start=1

/product="large tumor suppressor 1"

/protein_id="AADI6883.1"

/db_xref="GI:4324436"

/translation="VOHSINRKOSWKGSKESLVQRHPSLGNVYRSESPNSQADV
GPELSSGTAFAQAHPNSQQRVNPVPPQVRSVTPPPRGRTPPPRGTTPPPPSWE
PSSQTKRISGMIEYIISRISLSPVPGAWQEGYPPPLTISPMNPSSQAQRALSSVPVGR
QLIMOTSKFNTPGRPGVGGGSDFIHQNVPTGTVTRPQPPPYPLTPANGQS
SALOTGASAPSPFANGVQSMVFNRSNNHMLYNIIVPGLQTAWPOSSSAQAQS
PSGHEIPIWQPNIPVRSFNPLGRASHANSOSPATVTATPAIQQPKVMK
VLKPELQATALATHPSWMPQVOTQPTPESECTASSVPVIPPVEAPSVQGPVPPY
KHLHONSPVPVPEYSYKPKDEQSLPKEDDESEKSDGSDGKKEKQITTSPIVR
KNKQDERRESRIQISPSQAFKFFMEQHVNLKSHQRLHKKQLQLENNMRVGLSQD
ADQMRKMLCKQESNYIRLKRAMDKSMFVKITLIGIGAFGEVCLARKYDTKALYATM
TLRKDKVLLRNNVAVRDLIAAEDVNNVRLYYSFQDKNLMYFMVDYIPGDDMM
LITRKGLFENLARYIAELTCAVESVHKMGFTHRDKPNILIDRGHTKLTDFGGL
LITRKGLFENLARYIAELTCAVESVHKMGFTHRDKPNILIDRGHTKLTDFGGL

gene

CDS

TCGRWTHDSKYYOSGDHPDSDMSNCRGDRKLPKLLERRAAROHORCLAH
SLVGPNTIABEVLRTGYTQCLDNWSGVILCEMLVGQPPFLAQTPLETKMKVLIHQ
TSLHTPPQAKISPEASDLIIKCRGPDRLKNGADEIKAHFPFKTIDFSDDLQOQSA
SYIPKTHPTDTSNFDVDPDKLMSDGESEENI SDTSGWKYKHKHPHAFYEFTRR
FFDNGYYPYKPKPEYEHISQSEQOQSDDEDOHTSSDGNRLDLYV*

BASE COUNT 946 a 764 c 696 g 807 t

Query Match	Score 3213;	DB 10;	Length 3213;
Best Local Similarity	100.0%;	Pred. No. 0;	
Matches 3213;	Conservative 0;	Mismatches 0;	Indels 0; Gaps 0;
QY 1	GTCCACATTCATTAACCGAACAACAGCTGGAAGTTCTAAAGAGTCTCTAGTTCT	60	
DB 1	GTCCACATTCATTAACCGAACAACAGCTGGAAGTTCTAAAGAGTCTCTAGTTCT	60	
QY 61	CAGAGACAGGGCCATCTCTAGGAGAAATGTGGTTTCTGTTCTGAAAGCCCAACTCA	120	
DB 61	CAGAGACAGGGCCATCTCTAGGAGAAATGTGGTTTCTGTTCTGAAAGCCCAACTCA	120	
QY 121	CAGGCGGATGTAGGAGACCTCTGTCTGGATCCGCAATTCAGCAATTCCTCAAGCTAC	180	
DB 121	CAGGCGGATGTAGGAGACCTCTGTCTGGATCCGCAATTCAGCAATTCCTCAAGCTAC	180	
QY 181	CAAGCAATGGACAGAGAGTGAACCCGCCACACCACTCAAGTTAGGAGTGTACTCT	240	
DB 181	CAAGCAATGGACAGAGAGTGAACCCGCCACACCACTCAAGTTAGGAGTGTACTCT	240	
QY 241	CCACCACCTCCGAGAGGCGAGCCACCTCCCGGAGGACCACTCCCGCTCCCGCTCA	300	
DB 241	CCACCACCTCCGAGAGGCGAGCCACCTCCCGGAGGACCACTCCCGCTCCCGCTCA	300	
QY 301	TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTAACTCTCCGA	360	
DB 301	TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTAACTCTCCGA	360	
QY 361	ATCTCCCTCTTCCACTGGGGCGTGGCAGAGGGTACCTCCACCACTCTTACCCT	420	
DB 361	ATCTCCCTCTTCCACTGGGGCGTGGCAGAGGGTACCTCCACCACTCTTACCCT	420	
QY 421	TCCTCCCATGAATCCCTTAGCAGGCTCAGAGGGCAATAGTTCTCCAGTTGTAGA	480	
DB 421	TCCTCCCATGAATCCCTTAGCAGGCTCAGAGGGCAATAGTTCTCCAGTTGTAGA	480	
QY 481	CAACCCATCATCATGAGAGTACTAGCAATTTAACTTTACACAGGGGACCTGGAGTT	540	
DB 481	CAACCCATCATCATGAGAGTACTAGCAATTTAACTTTACACAGGGGACCTGGAGTT	540	
QY 541	CAGAAATGGTGGTGGTCTGATTTATCGTCACCAAAATGTCCCACTGGTCTGTG	600	
DB 541	CAGAAATGGTGGTGGTCTGATTTATCGTCACCAAAATGTCCCACTGGTCTGTG	600	
QY 601	ACTCGGAGCCACCACTCCATATCTCTGACCCAGCTAATGGCAAGGCCCTCTGCT	660	
DB 601	ACTCGGAGCCACCACTCCATATCTCTGACCCAGCTAATGGCAAGGCCCTCTGCT	660	
QY 661	TTACAAACAGGGCTTCTGCTGCTCCACCATCATTCGCCAATGGAACGTTCTCAGTCG	720	
DB 661	TTACAAACAGGGCTTCTGCTGCTCCACCATCATTCGCCAATGGAACGTTCTCAGTCG	720	
QY 721	ATGATGTGGCCCAACAGAGACAGTCAATAGAGCTTTTAAATTAATGTCCCTGGA	780	
DB 721	ATGATGTGGCCCAACAGAGACAGTCAATAGAGCTTTTAAATTAATGTCCCTGGA	780	
QY 781	CTGCAACAGCTGGGCCAGTCTGCTCTCTGCTCCGCGAGTCATCCCAAGCGTGGG	840	
DB 781	CTGCAACAGCTGGGCCAGTCTGCTCTCTGCTCCGCGAGTCATCCCAAGCGTGGG	840	
QY 841	CATGAATTCCTACATGGCAACCTTAACATACAGTGGAGTCAATTTTAAATACCCA	900	
DB 841	CATGAATTCCTACATGGCAACCTTAACATACAGTGGAGTCAATTTTAAATACCCA	900	
QY 901	TTAGGAAGTAGAGCAAGTCACTCTGCTTAATCTCAGCCTTCTGCCACTACAGTCACTGCC	960	

DB 901	TTAGGAAGTAGAGCAAGTCACTCTGCTTAATCTCAGCCTTCTGCCACTACAGTCACTGCC	960
QY 961	ATCACACCGGCTCTTATTAACAGAGCCCGTGAAGCATGCGCGTCTCAAAACAGAGCTG	1020
DB 961	ATCACACCGGCTCTTATTAACAGAGCCCGTGAAGCATGCGCGTCTCAAAACAGAGCTG	1020
QY 1021	CAGACTGCTTTAGCCCAACCCATCTTCTTGGATGCCACAGCCAGTTTCAGACTGTTCCAG	1080
DB 1021	CAGACTGCTTTAGCCCAACCCATCTTCTTGGATGCCACAGCCAGTTTCAGACTGTTCCAG	1080
QY 1081	CCTACCCCTTTTCTGAGGGTACAGTCTCAAGTGTGCGTCTCATCCACCTGTTCTGCTGAA	1140
DB 1081	CCTACCCCTTTTCTGAGGGTACAGTCTCAAGTGTGCGTCTCATCCACCTGTTCTGCTGAA	1140
QY 1141	GCTCCAGCTATCAAGTGTCCACCCGCTTATCCAAACATCTCTACACCAAAACCCA	1200
DB 1141	GCTCCAGCTATCAAGTGTCCACCCGCTTATCCAAACATCTCTACACCAAAACCCA	1200
QY 1201	TCGTGCTCCCTCCATATAGTCAAGTAAAGTAAAGCCCTCAAGATGAACAGCTTAGCTTACC	1260
DB 1201	TCGTGCTCCCTCCATATAGTCAAGTAAAGTAAAGCCCTCAAGATGAACAGCTTAGCTTACC	1260
QY 1261	AAGGAAGATGATGAGAGAGTGGGACAGTGGTACTCTGGGGATAAAGAAAGAAA	1320
DB 1261	AAGGAAGATGATGAGAGAGTGGGACAGTGGTACTCTGGGGATAAAGAAAGAAA	1320
QY 1321	CAGATTACAACCTTCACTATCTGTCGAAAAAACAAGAGATGAAGAACGAAGAG	1380
DB 1321	CAGATTACAACCTTCACTATCTGTCGAAAAAACAAGAGATGAAGAACGAAGAG	1380
QY 1381	TCCTGGATTCAGAGTACTTCCACAGGCTTTAAAGTCTTCATGGAGCAGCAGCTAGAG	1440
DB 1381	TCCTGGATTCAGAGTACTTCCACAGGCTTTAAAGTCTTCATGGAGCAGCAGCTAGAG	1440
QY 1441	ACGTCCTCAAGTCTCATCAGCAGGCTGTCATCGGAAGCAGCTAGAAAATGAAATG	1500
DB 1441	ACGTCCTCAAGTCTCATCAGCAGGCTGTCATCGGAAGCAGCTAGAAAATGAAATG	1500
QY 1501	ATCGGGTTGGATTTCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAA	1560
DB 1501	ATCGGGTTGGATTTCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAA	1560
QY 1561	GAGTCTAACTATATCTGTTTAAAGGGCTAAATGGCAAGCTATGTTCTTAAGATA	1620
DB 1561	GAGTCTAACTATATCTGTTTAAAGGGCTAAATGGCAAGCTATGTTCTTAAGATA	1620
QY 1621	AAGACATTAGGAATAGGAGGCTTTGGTGAAGTCTGTCTAGCAAGAAAGTCGATACTAAA	1680
DB 1621	AAGACATTAGGAATAGGAGGCTTTGGTGAAGTCTGTCTAGCAAGAAAGTCGATACTAAA	1680
QY 1681	GCTTTGTATGCAACAAGACTCTTCGAAAGAGAGCTTCTGCTCCGAAATCAGGTGGCT	1740
DB 1681	GCTTTGTATGCAACAAGACTCTTCGAAAGAGAGCTTCTGCTCCGAAATCAGGTGGCT	1740
QY 1741	CATGTCAGAGCGGAGAGGATATCTTAGCAGAGCCGCAATCAGTGGCTGGTCCGCTG	1800
DB 1741	CATGTCAGAGCGGAGAGGATATCTTAGCAGAGCCGCAATCAGTGGCTGGTCCGCTG	1800
QY 1801	TACTACTCTTCCAGGACAGGACAACTTGTACTTGTGATGGACTACATTCCTGGGGG	1860
DB 1801	TACTACTCTTCCAGGACAGGACAACTTGTACTTGTGATGGACTACATTCCTGGGGG	1860
QY 1861	GATATGATGAGCTTAAATAGAAATGGCACTTTTCTTGAATCTGCAGATTTCTAC	1920
DB 1861	GATATGATGAGCTTAAATAGAAATGGCACTTTTCTTGAATCTGCAGATTTCTAC	1920
QY 1921	ATAGCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGTGTATTCATAGAGAT	1980
DB 1921	ATAGCAGAACTTACCTGTGAGTGAAGTGTTCATAAAATGGTGTATTCATAGAGAT	1980
QY 1981	ATTAAACCTGATAACATTTGATGACCGTGTGGCCATTAATTAATGACTGACTTGGC	2040
DB 1981	ATTAAACCTGATAACATTTGATGACCGTGTGGCCATTAATTAATGACTGACTTGGC	2040

Db	1981	ATTTAAACCTGATTAACATTTTGTGATTGACCGCTGATGCCCATATTTAAATTGACTGACTGCTTTGGC	2041
Qy	2041	TTGTGCACTGGCTTCAGATGGACACATGACTCCAAAGTACTACCAGAGTGGGGATCACCCA	2101
Db	2041	TTGTGCACTGGCTTCAGATGGACACATGACTCCAAAGTACTACCAGAGTGGGGATCACCCA	2100
Qy	2101	CGGCAGATAGCATGGATTTCAGTAAACGAATGGGAGATCCTTCCAATTTGTCGGTGTGGG	2160
Db	2101	CGGCAGATAGCATGGATTTCAGTAAACGAATGGGAGATCCTTCCAATTTGTCGGTGTGGG	2160
Qy	2161	GACAGACTGAAGCCACTGGAGCGGAGAGCTGCTGCCAGCACCAGCGATGCTTAGCCCAT	2220
Db	2161	GACAGACTGAAGCCACTGGAGCGGAGAGCTGCTGCCAGCACCAGCGATGCTTAGCCCAT	2220
Qy	2221	TCCTCTGTTGGGACTCCCAATTATATTCACCTGAAGTGCTACTGCGAACAGAGATATACA	2280
Db	2221	TCCTCTGTTGGGACTCCCAATTATATTCACCTGAAGTGCTACTGCGAACAGAGATATACA	2280
Qy	2281	CAGCTGTGTGACGTGTGGAGTGTGTGTGTTATTCCTTTGTGAAATGTTGGTGGGACAACCT	2340
Db	2281	CAGCTGTGTGACGTGTGGAGTGTGTGTGTTATTCCTTTGTGAAATGTTGGTGGGACAACCT	2340
Qy	2341	CCCTTTCTTGGCACAAACCCCATATAGAAACACAAATGAAGGTTATCATCTGGCAAACTTCT	2400
Db	2341	CCCTTTCTTGGCACAAACCCCATATAGAAACACAAATGAAGGTTATCATCTGGCAAACTTCT	2400
Qy	2401	CTACACATCCCTCCTCAAGCTAAGCTGAGTCTGAAGCCTCTGAAGCCTCTGACCTCATTTATCAAACTG	2460
Db	2401	CTACACATCCCTCCTCAAGCTAAGCTGAGTCTGAAGCCTCTGACCTCATTTATCAAACTG	2460
Qy	2461	TGTCGAGGACCAAGAAGACCGCTCGGCAAGAACGGTGTGATGAGATATAAGGCTCATCCA	2520
Db	2461	TGTCGAGGACCAAGAAGACCGCTCGGCAAGAACGGTGTGATGAGATATAAGGCTCATCCA	2520
Qy	2521	TTTTTTTAAAGCAATCGATTTCTCTAGTGTATCTGAGACAGCAGTCTGCTCATACATCCCT	2580
Db	2521	TTTTTTTAAAGCAATCGATTTCTCTAGTGTATCTGAGACAGCAGTCTGCTCATACATCCCT	2580
Qy	2581	AAAAATCACGCATCCAAACAGATACATCCAAATTTGCAACCTGTGTGATCCTGATAAATTTGTG	2640
Db	2581	AAAAATCACGCATCCAAACAGATACATCCAAATTTGCAACCTGTGTGATCCTGATAAATTTGTG	2640
Qy	2641	AGCGATGGCAGGAGGAGGAAATATCAGTGACACTCTGACGGGATGGTATAAAAAATGGG	2700
Db	2641	AGCGATGGCAGGAGGAGGAAATATCAGTGACACTCTGACGGGATGGTATAAAAAATGGG	2700
Qy	2701	AAGCACCCGAGCAGCGCTTTCTATCAGTTTCACCTTTTCGGAGGTTTTTGTGATGACATGGC	2760
Db	2701	AAGCACCCGAGCAGCGCTTTCTATCAGTTTCACCTTTTCGGAGGTTTTTGTGATGACATGGC	2760
Qy	2761	TACCCATATAAATTTCCAAAGCCTATTACGATGATGAATACATTCATTACAGGCTCAGAA	2820
Db	2761	TACCCATATAAATTTCCAAAGCCTATTACGATGATGAATACATTCATTACAGGCTCAGAA	2820
Qy	2821	CAACAGCTGTGATGAAGATGATCAACACACAAAGCTCCGATGGAAACAAACCGAGATCTAGTG	2880
Db	2821	CAACAGCTGTGATGAAGATGATCAACACACAAAGCTCCGATGGAAACAAACCGAGATCTAGTG	2880
Qy	2881	TATGTTTTATAAATAGAGATCATTTGTAAGAAATTTGCAAGAGCCTGAAGTCAGGGGT	2940
Db	2881	TATGTTTTATAAATAGAGATCATTTGTAAGAAATTTGCAAGAGCCTGAAGTCAGGGGT	2940
Qy	2941	TTTTTGAAGTTTTGAGAAAATTTATGCAAAATGTGACAGAGTTTGTGCTGCTGCTACAATA	3000
Db	2941	TTTTTGAAGTTTTGAGAAAATTTATGCAAAATGTGACAGAGTTTGTGCTGCTGCTACAATA	3000
Qy	3001	TTTTTATTTTCCCTAAGTTATGGGAAATTTGTTTTTAAATTTGTTAATTTATTCACCCCTTTTAA	3060
Db	3001	TTTTTATTTTCCCTAAGTTATGGGAAATTTGTTTTTAAATTTGTTAATTTATTCACCCCTTTTAA	3060
Qy	3061	TTTCAGTAAATTTAGAAAAAATTTGTTTATAGGAAACTAAATTTATGACTGAGTATATAGTCT	3120
Db	3061	TTTCAGTAAATTTAGAAAAAATTTGTTTATAGGAAACTAAATTTATGACTGAGTATATAGTCT	3120

Qy	3121	AAATCTTGGTACTTAAAGTACTTTAAAGAGAGAGAGCTGGTATCTTTTGTATATATAATAA	3180
Db	3121	AAATCTTGGTACTTAAAGTACTTTAAAGAGAGAGCTGGTATCTTTTGTATATATAATAA	3180
Qy	3181	ATAATTTTAAATCCCAAAAAAAAAAAAAAAAAA	3213
Db	3181	ATAATTTTAAATCCCAAAAAAAAAAAAAAAAAA	3213
RESULT 3			
AR201455			
LOCUS			
DEFINITION			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
ORGANISM			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
FEATURES			
source			
BASE COUNT			
ORIGIN			
Query Match: 77.3%; Score 2483; DB 6; Length 3984;			
Best Local Similarity 87.4%; Pred. No. 0;			
Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps			
Qy	1	GTGCACATTCAAATTAAACGGAAACAAAGCTGGAAGGTTCTTAAAGAGTCTCTAGTTCCT	60
Db	732	GTGCAGCAATCAGTTAAACCGCAACAGAGCTGGAAGGTTCTTAAAGATCCTTAGTTCCT	791
Qy	61	CAGACAGACGGCCCATCTCTAGAGAGAAATGTGTTTATCGTCTGAAAGCCCAACTCA	120
Db	792	CAGAGCATGGCCCGCCACTAGAGAGAAAGTGTGGCTATCATCTCTGAGAGTCCCAACTCA	851
Qy	121	CAGCGGATGTAGGAAGACTCTGTCTGGATCCGCATTGCAGCATTTGCTCAAGCTCAC	180
Db	852	CAGACAGATGTAGAAAGACTTTGTCTGATCTGATATCAGCATTTGTTCAAGCTCAC	911
Qy	181	CCAAGCAATGGACAGAGAGTGAACCCCAACACCACTCAAGTAGGAGTGTACTTCCT	240
Db	912	CCTAGCAACGACAGAGAGTGAACCCCAACCACTCAAGTAGGAGTGTACTTCCT	971
Qy	241	CCACCACTCCGAGAGCCAGACCCCACTCCCGAGGACCACTCCCTCCCTCCCTCA	300
Db	972	CCACCACTCCCAAGAGCCAGACTCCCTCCCAAGAGGTCAACTCCACCTCCCTCTCA	1031
Qy	301	TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCCGA	360
Db	1032	TGGGAACCAACTCTCAACAAAGCGCTATTCTGGAACATGGATACGTAATCTCCCGA	1091
Qy	361	ATCTCCCTGTTCCACTCTGGGCGGTGGCAGGAGGGTACCTCCACCACTCTTACCAC	420
Db	1092	ATCTCTCTCTCCCACTGGGCGATGGCAAGAGGCTATCTCCACCACTCTCAACACT	1151
Qy	421	TCCTCCATGAATCCCTAGCCAGCTCAGAGGCCATTAGTCTGTGTTGAGTTGGTAGA	480
Db	1152	TCCTCCATGAATCCCTAATCAAGCAGAGAGAGCATTAGTCTGTGTTGTTGGCAGA	1211
Qy	481	CAACCCATCATATGCAGAGTACTAGCAAAATTAACCTTACACACGGCGACCTGGAGTT	540
Db	1212	CAACCAATCATATGCAGAGTCTAGCAAAATTAACCTTCCATCAGGAGACCTGGATG	1271
Qy	541	CAGATGTGGTGTAGTCTGATTTTATCGTGCACCAAAA---TGTCCTCCCACTGGTTCT	597
Db	1272	CAGATGTGTAGCAAACTGATTTATGATACACCAAAATGTGTCTCTCTGGCACT	1331

Qy 598 GTGACTCGGAGCCACCACCTCCATATATCTCTGACCCAGCTAATGGACAAAGCCCTCT 657
 Db 1332 GTGAATCGGAGCCACCACCTCCATATCTCTGACAGCAGCTAATGGACAAAGCCCTCT 1391
 Qy 658 GCTTTACAAACAGGGCTTCTGCTCCACCAATCGGCAATGGAAAGCTTCTCTCAG 717
 Db 1392 GCTTTACAAACAGGGGATCTGCTCTCTCGTCATATACAAATGGAAATATTCCTCAG 1451
 Qy 718 TCGATGATGGTGGCCCAACAGGAACAGTACATACATGGAGCTTTATAATTAATGTCCT 777
 Db 1452 TCTATGATGGTGGCCCAACAGGAATAGTCTATACATGGAACTATATAACATTTAGTACT 1511
 Qy 778 GGACTGCAAAACAGCGCTGGCCCGAGTCGTCTCTGCTCTGCGAGTCCATCCCCAAGCGT 837
 Db 1512 GGACTGCAAAACAAATTTGGCCCTCAGTCATCTCTGCTCTGCGAGCCAGTATCCCCGAGCT 1571
 Qy 838 GGGCATGAATTCCTACATGCGCAACCTAACATACCAGTGAGGTCAATCTTTTAATAC 897
 Db 1572 GGGCATGAATTCCTACATGCGCAACCTAACATACCAGTGAGGTCAATCTTTTAATAC 1631
 Qy 898 CCATTTAGGAAGTAGAGCAAGTCACTCTCTAATCTCAGCCTCTGCGCACTACAGTCACT 957
 Db 1632 CCATTTAGGAATAGAGCAAGTCACTCTCTAATCTCAGCCTCTGCGCACTACAGTCACT 1691
 Qy 958 GCCATCACACCCGCTCTTATTAACAGCCCGTGAAGCAATGCGCGTCTCTGAAACCCAGAG 1017
 Db 1692 GCAATTAACACAGCTCTTATTAACAGCCCGTGAAGCAATGCGGTGTTATTAACCCAGAG 1751
 Qy 1018 CTGCAGACTGCTTTAGCCCGCAACCCATCTCTTTGGATGCCACGCCAGTTCAGACTGT 1077
 Db 1752 CTACAGACTGCTTTAGCACCTTACACACCTCTTTGGATACCACGCCAATTCACAACTGT 1811
 Qy 1078 CAGCCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCTATCCACCTGTTGCT 1137
 Db 1812 CAACCCAGTCTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCTATCCACCTGTTGCT 1871
 Qy 1138 GAAGCTCCAGCTATCAAGGTCACCCAGCTTATCCAAACATCTGCTACACCAAAAC 1197
 Db 1872 GAAGCTCCAACTATCAAGGTCACCCAGCTTATCCAAACATCTGCTGCACCAAAAC 1931
 Qy 1198 CACTCTGCTCTCATATAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1257
 Db 1932 CACTCTGCTCTCATATAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAGTCAAG 1991
 Qy 1258 CCAAGGAAGATGATGAGTGAAGAGTGGGACAGTGGTCACTCTGGGGATAAAGAAAG 1317
 Db 1992 CCAAGGAAGATGAGTGAAGAGTGGTCACTCTGAGTGGGATTAAGAAAG 2051
 Qy 1318 AAACAGATTACACTTCACTATCACTGCTCGGAAACCAAGAAAGATGAAGAACGAAGA 1377
 Db 2052 AAACAGATTACACTTCACTATCACTGCTCGGAAACCAAGAAAGATGAAGAACGAAG 2111
 Qy 1378 GAGTCTCGATTACAGTCTACCTCCACAGCGCTTTAGTCTTCTCATGGACGACGCTA 1437
 Db 2112 GAATCTCGATTACAGTCTACCTCCACAGCGCTTTAGTCTTCTCATGGACCAACATGA 2171
 Qy 1438 GAGAACGTCCTGAAGTCTCATTACAGCAGCTGCTGATCGGAAGAGCAGCTAGAAAATGAA 1497
 Db 2172 GAATGTCCTCAATCTCATTACAGCAGCTGCTGATCGTAAGAAACATTAAGAAATGAA 2231
 Qy 1498 ATGATGCGGGTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAG 1557
 Db 2232 ATGATGCGGGTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAA 2291
 Qy 1558 AAAGAGTCTACTATATCTCTTAAAGGCTTAAATGGACAGTCTATGTTTGTAAAG 1617
 Db 2292 AAAGATCTAAATCTACCTCGCTTAAAGGCTTAAATGGACAGTCTATGTTTGTAAAG 2351
 Qy 1618 ATAAGACATTAGGAATAGGAGCTTTGGTGAAGTCTGCTAGCAAGAAAGTCAAGTACT 1677
 Db 2352 ATAAGACATTAGGAATAGGAGCTTTGGTGAAGTCTGCTAGCAAGAAAGTCAAGTACT 2411

Qy 1678 AAAGCTTTGTATGCAACAAAGACTTTCGAAGAAAGAGCTTCTGCTCCGAAATCAGGTG 1737
 Db 2412 AAAGCTTTGTATGCAACAAAGACTTTCGAAGAAAGAGTCTTCTTCGAAATCAAATTC 2471
 Qy 1738 GCTCATGTGAAGCGGAGAGGATATCTAGCAGAACCGCAATGAGTGGTGGTCTCCG 1797
 Db 2472 GCTCATGTGAAGCGGAGAGGATATCTGCTGAAGTGAACATGAATGGTGGTCTCGT 2531
 Qy 1798 CTGTACTACTCTTCCAGGACAAAGCAACTTGTACTTTGTGATGGACTACATCTCTGGG 1857
 Db 2532 CTATATATTATCTCAAGATGAAGCAATTTATATTTGTAATGGACTACATCTCTGGG 2591
 Qy 1858 GGGATATGATGAGCCTATTATTAAGATGGCAGTCTTCTCTGAAATCTGGCAGATTC 1917
 Db 2592 GGTGATATGATGAGCCTATTATTAAGATGGCAGTCTTCTCTGAAAGTCTGGCAGATTC 2651
 Qy 1918 TACATAGCAGAACTTACCCTGCGAGTTCAGAACTCTTCAAAATGGGTTTATTCATAGA 1977
 Db 2652 TACATAGCAGAACTTACCCTGCGAGTTCAGAACTCTTCAAAATGGGTTTATTCATAGA 2711
 Qy 1978 GATATTAACCTGATATATTTGATTCAGCTGATGGCCATATTAATTTGACTGACTTT 2037
 Db 2712 GATATTAACCTGATATATTTGATTCAGCTGATGGCCATATTAATTTGACTGACTTT 2771
 Qy 2038 GGTGTGCTCACTGGCTTCAGATGAGACATGACTCCAAAGTACTACAGAGTGGGATCAC 2097
 Db 2772 GGCCTCTCACTGGCTTCAGATGAGACATGACTCCAAAGTACTACAGAGTGGGATCAC 2831
 Qy 2098 CCAGGCAAGATAGCATGATTTTCCAGTAAACAAATGGGAGATCTTCCAAATTTGGGTG 2157
 Db 2832 CCAGGCAAGATAGCATGATTTTCCAGTAAACAAATGGGAGATCTTCCAAATTTGGGTG 2891
 Qy 2158 GGGGACAGACTGAAGCCACTGGAGGAGAGTCTCTCCAGCAGCAGCAGTGTCTAGCC 2217
 Db 2892 GGAGACAGACTGAAGCCACTGGAGGAGAGTCTCTCCAGCAGCAGCAGTGTCTAGCC 2951
 Qy 2218 CATTCTCTGTTGGGACTCCCAATATATGCACTGAAAGTCTACTCGCAACAGGATAT 2277
 Db 2952 CATTCTCTGTTGGGACTCCCAATATATGCACTGAAAGTCTACTCGCAACAGGATAT 3011
 Qy 2278 ACACAGCTGTGAGTGGTGGAGTGTGCTGTTATCTTCTGAAATCTTGGTGGGACAA 2337
 Db 3012 ACACAGTGTGAGTGGTGGAGTGTGCTGTTATCTTCTGAAATCTTGGTGGGACAA 3071
 Qy 2338 CCTCTCTTCTTGGGCAAAACCCCTTATGAACACAAATGAAGTTATCATCTGGCAAACT 2397
 Db 3072 CCTCTCTTCTTGGGCAAAACCCCTTATGAACACAAATGAAGTTATCATCTGGCAAACT 3131
 Qy 2398 TCTCTACACATCCCTCTCAAGCTAAGCTGAGTCTGAGGCTCTGAGCTCTATTCATA 2457
 Db 3132 TCTCTACACATCCCTCTCAAGCTAAGCTGAGTCTGAGGCTCTGAGCTCTATTCATA 3191
 Qy 2458 CTGTCTCAGCAGCAGAGAGCCGCTCGCAAGAGAGCTGCTGATGAGATGAAGGCTCAT 2517
 Db 3192 CTGTCTCAGCAGCAGAGAGCCGCTCGCAAGAGAGCTGCTGATGAGATGAAGGCTCAT 3251
 Qy 2518 CCATTTTAAAGACCATCGATTTCTTCTAGTGTCTGAGCAGCAGCTCTGCTTCATACATC 2577
 Db 3252 CCATTTTAAAGACCATCGATTTCTTCTAGTGTCTGAGCAGCAGCTCTGCTTCATACATC 3311
 Qy 2578 CCTAAATCAGCATCCAAACAGATACATCCAAATTTGAGCCCTGTTGATCTCTGATAAATG 2637
 Db 3312 CCTAAATCAGCATCCAAACAGATACATCCAAATTTGAGCCCTGTTGATCTCTGATAAATG 3371
 Qy 2638 TGGAGCATGCGCAGGAGGAGAAATATCAGTCACTCTGAGGAGTGTGATAAAT 2697
 Db 3372 TGGAGCATGCGCAGGAGGAGAAATATCAGTCACTCTGAGGAGTGTGATAAAT 3431
 Qy 2698 GGGAGCACCCTGAGCAGCTTCTTATGAGTTCACCTTTTCGGAGGTTTTTTCATGACAAAT 2757
 Db 3432 GGGAGCACCCTGAGCAGCTTCTTATGAGTTCACCTTTTTCGGAGGTTTTTTCATGACAAAT 3491
 Qy 2758 GGCTACCCATATATATTCACAAAGCCTATTCAGTATGATGATACATTCATTCACAGGCTCA 2817

Db	2670	ATAAGACACTAGGAATAGGAGCAATTTGGTGAAGTCTGTCTAGCAAGAAAAAGTAGACT	27272
Qy	1678	AAAGCTTTGTATGCAACAAAGACTCTTGCAGAAAGAGCTTCTGCTCCGAAATACAGTG	17373
Db	2730	AAGCTTTGTATGCACAAAAACTCTTCGAAAGAAGATGTTCTTTCGAAATCAAGTC	27899
Qy	1738	GCTCATGTGAAGCGGAGGGGATATCCTTAGCAGAAGCCGACAAATGAGTGGGTGGTCCGC	17977
Db	2750	GCTCATGTAAGGCTGAGAGAGATATCCTGGCTGAAGCTGACAATGAATGGTAGTTCGT	2849
Qy	1798	CTGTACTACTCTTTTCAGGACAAGGACAACACTTGTACTTTGTGATGGCATACATTCCTGGG	1857
Db	2850	CTATATTATTCATTCCAAGATAAGGACAATTTATACTTTGTAATGGACTACATTCCTGGG	2909
Qy	1858	GGGATATGATGAGCCTATTAAATTCGAATGGGCATCTTTCCTGAAAAATCTGSCAGATTC	1917
Db	2910	GGTGATATGATGAGCCTATTAAATGAGATGGGCATCTTTCGAAAAAGTCTGGCAGATTC	2969
Qy	1918	TACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGA	1977
Db	2970	TACATAGCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGA	3029
Qy	1978	GATATTAAACCTGATACAACTTTTGTATGACCTGATGGCCATATTAATGACTGACTTT	2037
Db	3030	GATATTAAACCTGATATATTTTGTATGATCGTATGGTCATATTAATTAATGACTGACTTT	3089
Qy	2038	GGCTTGTGCACCTGGCTTCAGATGGACACATGACTCCAAGTACTACCAGAGTGGGATCAC	2097
Db	3090	GGCCTGTGCACCTGGCTTCAGATGGACACACGATTTCTAAGTACTATCAGAGTGGTGACCAT	3149
Qy	2098	CCAGCGCAAGATAGCATGGATTTCACTAACGAATGGGAGATPCTTCCAATTTGTCGGTGT	2157
Db	3150	CCAGCGCAAGATAGCATGGATTTCACTAATGAATGGGGGATCCCTCAAGCTGTCTGATGT	3209
Qy	2158	GGGACAGACTGAAGCCACTGGAGCGGAGAGCTGTGCCAGCACACGAGATGTCTAGCC	2217
Db	3210	GGAGACAGACTGAAGCCACTTAGAGCGGAGAGCTGCAGCGCAGCACGAGATGTCTAGCA	3269
Qy	2218	CATTCTCTGGTTGGGACTCCCAATATATTTGCACCTGAAGTGTCTACTCGGAACAGGATAT	2277
Db	3270	CATTCTTTGGTTGGGACTCCCAATATATTTGCACCTGAAGTGTTCGTACGAAACAGGATAC	3329
Qy	2278	ACACAGCTGTGTGACTGGTGGAGTGTGGTGTATTCTTTTGTGAATGTCTGGTGGGACAA	2337
Db	3330	ACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTCTTTTGTGAATGTCTGGTGGGACAA	3389
Qy	2338	CCTCCTTTCTGGGACAAACCCCATTTAGAAAACAAATGAAGGTTATCATCTGCGCAAACT	2397
Db	3390	CCTCCTTTCTGGGACAAACCCCATTTAGAAAACAAATGAAGGTTATCAACTGGCAAACA	3449
Qy	2398	TCCTACACATCCTCTCAAGCTAAGCTGACTCCTGAAGCCTCTGAGCTCATATTATCAAA	2457
Db	3450	TCCTCTCACATTTCCACCAACAGCTAAACTCACTGCTGAAGCTTCTGTATTATTATTA	3509
Qy	2458	CTGTGTCGAGGACGAGAAGACCGCCTCGGCAAGAAGCGTGTGATGAGATAAAGGCTCAT	2517
Db	3510	CTTTGGCGAGGCCGCAAGATCGCTTAGCACAAGATGGTGTGATGAATAAAGGCTCAT	3569
Qy	2518	CCATTTTTTAAGACCATCGATTTCTTAGTGATGTGAGACAGCAGCTGTGCTCATACATC	2577
Db	3570	CCATTTTTTAAAACAATTGACTTCTCCAGTGCACCTGAGACAGCAGCTGTGCTCATACATT	3629
Qy	2578	CCTAAAAATCAGCATCCAACAGATACATCCAATTTTGACCCCTGTTGATCCTGATAAATTG	2637
Db	3630	CCTAAAAATCACACCCCAACAGATACATCAAAATTTTGATCCTGTTGATPCTGATAAAAATTA	3689
Qy	2638	TGGAGCATGGCAGCGAGGAGGAAAAATATCACTGACACTCTGACGCGGATGCTATAAAAAT	2697
Db	3690	TGGAGTGATGATACAGAGGAGNAAATGTAATGACACTCTCAATGGATGATATAAAT	3749
Qy	2698	GGGAAGCACCCCGAGCACGCTTCTTATGAGTTCACTCTTCGGAGGTTTTTTTATGACAAAT	2757
Db	3750	GGAAAGCATCCTGCAACATGATCTATGAATTTTACTCTCCGAAGGTTTTTTTATGACAAAT	3809

```

OY 2758 GGCTACCCATATAATTAATCAAGGCTATTGAGTATGAATACATCAATTCACAGGCTCA 2817
Db 3810 GGCTACCCATATAATTAATCAAGGCTATTGAGTATGAATACATCAATTCACAGGCTCA 3869
OY 2818 GAACACAGCTCTGATCAAGATGATCAACACACAGCTCCGATGGAAACACAGGAGATCTA 2877
Db 3870 GAGCAGAGCTCGGATGAGATGATCAAAACACAGGCTCAGAGATTAAAAATCGCATCTA 3929
OY 2878 GTGTATGTTTAACTAGGAGATCATTTGTA---GAATTTGCAAGAGGCTGAAGTG 2933
Db 3930 GTATATGTTTACACACATAGTAATAATGTAATGAGGATTTGTAAGAGGCTGAATG 3989
OY 2934 CAGGGGTTTGAAGTTTGTAG---AAATTTGCAAAATGTGACAGAG-----TTTGTG 2985
Db 3990 CGAGGTGTTTGTAGGTTCTGAGAGTAAATTTATGCAAAATATGACAGAGCTATATGCTG 4049
OY 2986 GCTCTGTGTACATATTTATTTCTTCTAGTTATGGGAAATGTTTAAATTTAAATTT 3045
Db 4050 GCTCTGTGTACATATTTATTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 4109
OY 3046 ATTCACCCCTTTTAACTAGTAAATTTAGAAAAATTTGTTATAGGAAAGTAAATATGAA 3105
Db 4110 ATTCACCCGTTTAACTAGTAAATTTAGAAAAATTTGTTATAGGAAAGTAAATATGAA 4169
OY 3106 CTGATATATAGTCAATTTCTTGGTACTTAAAGTACTTAAAGAGAGAG 3154
Db 4170 CTGATATATAGTCAATTTCTTGGTACTTAAAGTACTTAAAGTAAAGTAG 4218

RESULT 5
AF164041 4241 bp mRNA linear PRI 06-DEC-1999
LOCUS Homo sapiens WARTS protein kinase (WARTS) mRNA, complete cds.
DEFINITION AF164041
ACCESSION AF164041
VERSION AF164041.1 GI:5738135
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 4241)
Kishimoto,T., Nawa,S.-I., Nagamine,Y., Nishiyama,Y. and Saya,H.
WARTS protein, polynucleotide encoding the same, antisense
polynucleotide thereof, and antibody recognizing the protein
Patent: Japan (PCT/JP98/03739) 24-AUG-1998;
Sumitomo Electric Industries, Ltd.; 1 Taya-cho, Sakae-ku,
Yokohama-shi; Kanagawa;
Japan.;
2. (bases 1 to 4241)
Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Yamamoto,H., Hiraoka,T., Kitamura,N. and
Saya,H.
A human homolog of Drosophila warts tumor suppressor, h-warts,
localized to mitotic apparatus and specifically phosphorylated
during mitosis
FEBS Lett. 459 (2), 159-165 (1999)
99447636
PUBMED 10518011
3 (bases 1 to 4241)
Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Nakamura,H., Koga,H. and Saya,H.
Direct Submission
Submitted (30-JUN-1999) Tumor Genetics and Biology, Kumamoto
University School of Medicine, 2-2-1 Honjo, Kumamoto 860-0811,
Japan
FEATURES
source
1. .4241
/organism="Homo sapiens"
/db_xref="taxon:9606"
1. .4241
/gene="WARTS"
33. .3425

```

```

/gene="WARTS"
/Note="similar to Drosophila warts tumor suppressor;
localized on mitotic apparatus, specifically
phosphorylated at mitotic checkpoint"
/codon_start=1
/product="WARTS protein kinase"
/protein_id="AAD50272.1"
/db_xref="GI:5738136"
/translation="MKRSEKPEGYRMPKTPPASNTYTVSSROMLOETRESLRLNLSKP
SDAAKAENMSKSTEDPRQVNPFPKFTGHKALQEIIRNSLLPFANETSSRSSTSEVN
POMLODQAAGFDEDMVLOALOKTNRSIEAAIEFISKMSYQDPRRQMAAAARPIN
ASMKPGNVQSQNRKQSKESLVPQRHGPGLGESVAYHSESPNSDVGRLPGLS
GISAFVQAHPSNGVRNPVPQVRSVTPPPRGTTPPPRGTTPPPRGTTPPPRGTTP
SSKNEVVISISVPVCAQWQEGVPPPLNTSPMNPNGQSGISSVPGVQPLIMOS
SSKNFSGRPMQNGTQDFTMIHQNVVAGTVNRPVPPYPLTAANGQSPSALQTS
GSAAPSYTNGSIFQSMVMVNRNHNELYNISVGLQTNWPSQSSAPSSSGHE
ITWOPNIPVRSNFPNPLGNRASHANSQATVTTAITPAIQOPVYKSMRVLKPEL
QTALAPHPISWIPQIQVQPPFEGTASNTVMPVAEAPNTQPPVPPYKRLHLO
NPSVPPYESISKPSKEDQSLPKEDSEKSYENDVSGDEKQKQITSPITVRKNKDE
ERRSRQISYSPQAFKPFMEHQVENVLKSHOORLHRKQLENNMRVGLSDQADQMR
KMLQCKESNYIRLAKRMDKSMFKVKTGLGIGAFGEVCLARKVDTKALYATKLRKD
VLLRNQVAVKAEKDILAEADNEWVRLYISFQDKNLIYFVMDIYIPGGDMSLIRMG
IFPESLARFYIAELTCAVESVHKMGFHRDILKPDNILDROGHKLKLDGFLGCTGRWT
HDSYVQSGDHPRODSMDFSNEWDGPPSCRCGRDLKPLERRAQRHQRLAHSVLGP
NYIAPEVLLRTGYQLCDWMSVGLFEMLVGQPPFLAQTPLETQVQVQVQVQVQVQV
POAKLSPEASDLIIKLCRGPEDLCKGADIEKAHPFKTIDFSSDLRQOASVPIKI
THPTDTSNFDVPDKLWSDNDEENVDNLNGYKNGKHPEHAFYETFRFFDDNG
YPYNPYPIEYIYINSQSEQSDDEDDOONTGSEIKNRDLVTY"
/misc_feature
4245. .3059
/gene="WARTS"
/Note="Region: serine/threonine kinase domain"
BASE COUNT 1384 a 884 c 817 g 1156 t
ORIGIN
Query Match 76.9%; Score 2472; DB 9; Length 4241;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 385; Indels 16; Gaps 5;
OY 1 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGGTTCTTAAAGAGTCTTACTCTCT 60
Db 534 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGGTTCTTAAAGAGTCTTACTCTCT 593
OY 61 CAGAGACACGGCCCATCTCTAGGAGAAATATGTTTCTGTTCTGAAAGCCCAACTCA 120
Db 594 CAGAGACATGCGCCCATCTAGGAGAAATATGTTTCTGTTCTGAAAGCCCAACTCA 653
OY 121 CAGCGGATGTAGGAGACCTCTGTCTGATCCGCGATTCAGCATTTGCTCAAGCTCAC 180
Db 654 CAGAGATGTAGGAGACCTTTGTCTGATCTGATATATCAGCATTTGTTCAAGCTCAC 713
OY 181 CCAAGCAATGGACAGAGTGAACCCGCCACCCACCTCAAGTTAGGAGTGTACTCTCT 240
Db 714 CAGCAACAGGACAGAGTGAACCCGCCACCCACCTCAAGTTAGGAGTGTACTCTCT 773
OY 241 CCACACCTCGGAGAGGCGCAGACCCACCTCCCGAGGACCCACCTCCCGCTCCCTTCA 300
Db 774 CCACACCTCGGAGAGGCGCAGACCCACCTCCCGAGGACCCACCTCCCGCTCCCTTCA 833
OY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGACACATGGAGTAGTATCTCCCGA 360
Db 834 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGACACATGGAGTAGTATCTCCCGA 893
OY 361 ATCTCCCTCTGTCACCTGGGCGTGGCAGAGGGGTACCTCCACCATCTTACCATC 420
Db 894 ATCTCTCTCTGTCACCTGGGCGTGGCAGAGGGGTATCTCTCCACCATCTTACCATC 953
OY 421 TCTCCCATGAATCCCTCTAGCAGGCTCAGAGGGGCATAGTTCTGTCTCCAGTGGTAGA 480
Db 954 TCCCCCATGAATCTCTCTAATCAGGACAGAGAGCATAGTTCTGTCTCTGTTGGCAGA 1013
OY 481 CAACCATCATCATGACAGAGTACTAGCAAAATTTAACTTTACACAGGCGACCTGGAGTT 540
Db 1014 CAACCATCATCATGACAGAGTACTAGCAAAATTTAACTTTTCCATCATGAGGAGACCTGGAATG 1073

```

Qy	1618	ATAAAGACATTAGGAATAGGACGGTTTGGTGAAGTCTCTCTAGCAAGAAAGATCGATACT	167
Db	2154	ATAAAGACACTAGGAATAGGACGATTTGGTGAAGTCTCTCTAGCAAGAAAGATGATACT	2213
Qy	1678	AAAGCTTTGTATGCACAAAGACTCTTGCAGAAAGAACGTTCTGCTCCGAAATACAGGTG	1737
Db	2214	AAGCGTTTGTATGCACAAAAAACCTCTCGAAGAAAGATGTTCTTCTCGAATCAAGTC	2273
Qy	1738	GCTCATGTGAAAGCGGAGAGGATATCTCTAGCAGAAGCCGACANATGAGTGGGTGGTCCGC	1797
Db	2274	GCTCATGTGTAAGGCTGAGAGAGATATCTGGCTGAAGCTGACAATGAATGGGTAGTTGCT	2333
Qy	1798	CTGTACTACTCTTTCCAGCACAAAGCAACTTGTACTTTGTGATGGACTACATTCCTGGG	1857
Db	2334	CTATATTATTCATCTCAAGATGAGGACAATTTATACATTGTGTAATGGACTACATTCCTGGG	2393
Qy	1858	GGGGATATGATGAGCCTATTAAATTAGAATGGGCATCTTTCTGAAAAATCTGGCACAGATTC	1917
Db	2394	GGTGATATGATGAGCCTATTAAATTAGAATGGGCATCTTTCCAGAAAAGCTCTGGCACGATTC	2453
Qy	1918	TACATAGCAGAACTTACCTGTCGCATCTGAAGTGTTCATAAAATGGCTTTTATTCATAGA	1977
Db	2454	TACATAGCAGAACTTACCTGTCGCATCTGAAGTGTTCATAAAATGGGTTTTTATTCATAGA	2513
Qy	1978	GATATTAAACCTGATAAACAATTTTGATTCACCGCTGATGCCATATTAAATTCAGTGACTTT	2037
Db	2514	GATATTAAACCTGATAAATATTTTGATTCATCGTGATGCTCATATTAATTCAGTGACTTT	2573
Qy	2038	GGCTTGTGCAGCTGCTTCAGATGGACACATGACTCCAAGTACTACAGAGTGGGGATCAC	2097
Db	2574	GGCCTCTGCAGCTGGCTTCAGATGGACACACGATTCCTAAGTACTATCAGAGTGGTGACCAT	2633
Qy	2098	CCAGCGCAAGATAGCATGGATTTTCAGTAAGAAATGGGGAGATCCTCCAAATTTGCGGTGT	2157
Db	2634	CCAGCGCAAGATAGCATGGATTTTCAGTAATGATGGGGGAGTCCCTCAAGCTGTCATGT	2693
Qy	2158	GGGGACAGACTGAAGCCACTGGAGCGGAGAGCTGCTGCCAGCACACGAGCATGCTAGCC	2217
Db	2694	GGAGACAGACTGAAGCCCAATTAGACGGGAGAGCTGCAGCCAGCACACGAGCATGCTAGCA	2753
Qy	2218	CATTCTCGTTGGGACTCCCAATTAATTGCACCTGAAGTGTCTACGCGAAACAGGATAT	2277
Db	2754	CATTCTTGGTTGGGACTCCCAATTAATTGCACCTGAAGTGTCTACGCGAAACAGGATAC	2813
Qy	2278	ACACGCTGTGTGACTGGTGGAGTCTGGTGTATTCTTTGTGAAATGTTGGTGGGACAA	2337
Db	2814	ACACAGTTGTGTGANTGGTGGAGTCTGGTGTATTCTTTTGAATATGTTGGTGGGACAA	2873
Qy	2338	CTCTCTTCTTGGCACAAACCCATTAGAAACACAAATGAGGTTATCATCTCGCAAACT	2397
Db	2874	CTCTCTTCTTGGCACAAACCCATTAGAAACACAAATGAGGTTATCATCTCGCAAACT	2933
Qy	2398	TCCTCTACACATCCCTCTCAAGCTAAGCTGAGTCTCTGAAGCCTCTGACCTCATTTATCAA	2457
Db	2934	TCCTCTTACATTCACCAACAGCTAACTCAGTCTCTGAGCTTCTGATCTTATTATTAA	2993
Qy	2458	CTGTGTGAGGACAGAGACGCCCTCGGCAAGAACGGTGCCTGATGAGATAAAGGCTCAT	2517
Db	2994	CTTTGCGGAGGACCGGAGATCGCTTTAGGCAAGAATGGTGTGATGAAATATAAAGCTCAT	3053
Qy	2518	CCATTTTTTAAGACGATCGATTTCCTAGTGCATCTGAGACAGAGCTGCTGCTCATACATC	2577
Db	3054	CCATTTTTTAAGACGATCGATTTCCTAGTGCATCTGAGACAGAGCTGCTGCTCATACATC	3113
Qy	2578	CCTAAATCAGCATCCCAACAGATACATCCAAATTTTCGACCCCTGTTGATCCTGATAAATTTG	2637
Db	3114	CCTAAATCAGCATCCCAACAGATACATCCAAATTTTCGATCCTGTTGATCCTGATAAATTA	3173
Qy	2638	TGAGCGGATGGCAGCGAGGAGGAATATCAGTGACACTCTGACCGGATGGTATAAAAT	2697
Db	3174	TGGAGTGATGTAACGAGGAGAAATGTAATGACACTCTCAATGGATGGTATAAAAT	3233
Qy	2698	GGGAAGCACCAGGACAGCGCTTCTATCAGTTTCACCTTTCGGAGGTTTTTTTGATGACAAT	2757

```

Db 3234 GGAAGCATCTGGAATGATCTTATGATTTACCTCCGAAGGTTTTTGTGATCAAT 3293
QY 2758 GGTACCCATATAATATCAAGCCCTATTGAGTATGAATACATTCATTCACAGGCTCA 2817
Db 3294 GGTACCCATATAATATCAAGCCCTATTGAGTATGAATACATTCATTCACAGGCTCA 3353
QY 2818 GAACAACAGTCTGATGAAGATGATCAACACAGAGTCCGATGGAACACCGAGATCTA 2877
Db 3354 GAGCAGCAGTCCGATCAAGATGATCAACACAGAGCTCAGAGATTAAAAATCGCATCTA 3413
QY 2878 GTGTATGTTTATAAATACTAGGAGATCATTTGTA --- GAATTTGCAAGAGCCCTGAAGTG 2933
Db 3414 GTATATGTTTAAACACACTAGTAAATAAATGATATGAGGATTTGTAAGAGGCCCTGAATG 3473
QY 2934 CAGGGTGTGTAAGTTTGTAG --- AAAATATGCAAAATGTGACAG --- TTTGTGT 2985
Db 3474 CGAGGTGTTTGTAGGTTCTGAGAGTAAATATGCAAAATGATGACAGCTATATGTGT 3533
QY 2986 GCTCTGTGACATATTTTATTTCTTAAGTTATGGAATTTGTTTAAATGTTAAATTT 3045
Db 3534 GCTCTGTGACATATTTTATTTCTTAAGTTATGGAATTTGTTTAAATGTTAAATTT 3593
QY 3046 ATT-CCACCTTTTATTCAGTAAATTTAGAAAAATTTTATTAAGAAAAAGTAAATATGA 3104
Db 3594 ATTCCAGCGCTTTAAATCAGTATTTAGAAAAAATTTGTTATAGGAAAAAGTAAATATGA 3653
QY 3105 ACTGAGTATATAGTCAATCTTGTGACTTAAAGTACTTAAAGAGAGAAG 3154
Db 3654 ACTGAATATATAGTCAGTTCTGGTACTTAAAGTACTTAAAGTAAAGTAG 3703

```

RESULT 6

E24613 2442 bp DNA linear PAT 18-JUN-2001

DEFINITION warts Protein, polynucleotide encoding the protein, antisense polynucleotide thereof and antibody recognizing the protein.

ACCESSION E24613.1 GI:13024646

VERSION JP 1999089580-A/1.

KEYWORDS unidentified.

SOURCE unidentified.

ORGANISM unclassified.

REFERENCE 1 (bases 1 to 2442)

AUTHORS Toshihiko, K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.

TITLE warts Protein, polynucleotide encoding the protein, antisense polynucleotide thereof and antibody recognizing the protein

JOURNAL Patent: JP 1999089580-A 1 06-APR-1999;

COMMENT SUMITOMO ELECTRIC IND LTD

OS Unidentified

PN JP 1999089580-A/1

PD 06-APR-1999

PF 24-SEP-1997 JP 1997258689

PR TOSHIOHKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI

YASUYUKI NISHIYAMA,

PI HIDEYUKI SAVA,

PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC

G01N33/532//C12P21/02, (C12P21/02, C12R1:19), C12N15/00,

PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,

CC Strandedness: Double;

CC Topology: Linear;

PH Key

FT source 1. 2442 Location/Qualifiers

FT /organism='Unidentified'.

1. 2442 Location/Qualifiers

/organism='unidentified'

/db_xref='taxon:32644'

777 a 528 c 491 g 646 t

BASE COUNT

ORIGIN

```

Query Match 60.6%; Score 1946.8; DB 6; Length 2442;
Best Local Similarity 87.9%; Pred. No. 0;
Matches 2146; Conservative 0; Mismatches 292; Indels 4; Gaps 2;
QY 331 TCTGGGAACATGAGTAGTAAATCTCCGAATATCTCCCTGTTCCACCTGGGGGTGGCAG 390
Db 1 TCTGGGAACATGAGTAGTAAATCTCCGAATATCTCCCTGTTCCACCTGGGGGTGGCAG 60
QY 391 GAGGGGTACCTCCACACCTCTTACCACCTTCTCCCATGAATCCCTAGCCAGGCTCAG 450
Db 61 GAGGGGTATCTCTCCACACCTTTTCAACACTTCCCTCATGAATCTCTTATCAAGACAG 120
QY 451 AGGGCCATTAGTTCTCTCCAGTTGGTAGACAAACCATCATCATCATGAGAGTACTAGCAA 510
Db 121 AGAGGCATTAGTTCTCTCTGTTGGCAGACAAACCAATCATCATGAGAGTTCTAGCAA 180
QY 511 TTAACTTTACACCGGCGACCTGGAGTTGAGATTTGAGATTTGAGTCTGATTTATTC 570
Db 181 TTAACTTTCCATFCAGGGAGACCTGGAATGCAAGTGGTACTGGACAAACTGATTTCTATG 240
QY 571 GTGCACCAAAA---TGTCCTCCACTGTTCTGTGACTCGGCAGCCACCATCTCCATATCCT 627
Db 241 ATACACCAAAATGTTGTCCTGCTGCACTGTGAATCGGCAGCCACCATCTCCATATCCT 300
QY 628 CTGACCCCTAGCTAATGACAAAAGCCCTCTGCTTTTACAAACAGGGCTTCTGCTGTCCA 687
Db 301 CTGACAGCAGCTAATGACAAAAGCCCTCTGCTTTTACAAACAGGGGATCTGCTGTCTCT 360
QY 688 CCATATTTGCGCAATGGAACGTTCTCAGTTCGATGATGGTCCCAACAGGAAACAGTCTAT 747
Db 361 TCGTCATATACAAATGGAAGTATTTCTCAGTGTATGATGGTCCCAACAGAAATAGTCTAT 420
QY 748 ACATGAGCTTTTAAATATATGTTCCCTGGACTCCAAACAGCCCTGGCCCCAGTCTCT 807
Db 421 AGATGGAATATATTAACATTTAGTGTGACTGGAATTCGAAACAAATTTGGGCTAGTCTAT 480
QY 808 TCTGCTCTCGCAGTCAATCCCAAGCGTGGGCATGAAATTTCTACATGGCAACCTAAC 867
Db 481 TCTGCTCCAGCCAGTCAATCCCGAGCAGTGGCATGAAATCCTACATGGCAACCTAAC 540
QY 868 ATACAGTGGGTCAAATTTCTTTTATAACCCATTTAGGAAGTAGAGCAAGTCACTCTGCT 927
Db 541 ATACAGTGGGTCAAATTTCTTTTATAACCCATTTAGGAAGTAGAGCAAGTCACTCTGCT 600
QY 928 AATTCACGCTTCTGCCACTACAGTCACTGCCATCACACCCCTCTTATTCACAGCCC 987
Db 601 AATTCACGCTTCTGCTACACAGTCACTGCAATTTACACCCAGTCTCTTATTCACAGCCT 650
QY 988 GTGAAAGCATGCGCTCTCTGAAACCCAGAGCTGACAGCTGCTTTAGCCCCCAACCCCTCT 1047
Db 661 GTGAAAGTATGCGTCTATTAAACCCAGAGCTACAGACTGCTTTAGCACCTACACACCT 720
QY 1048 TCTTGTATGCCAGCCAGTTCAGACTGTTACGCTTACCCCTTACCCCTTTTCTGAGGTTAGCT 1107
Db 721 TCTTGTATGCCAGCCAAATTTCAACTGTTTCAACCCAGTCTCTTCTGAGGAAACCGCT 780
QY 1108 TCAAGTGTGCTCTCATCCACCTGTTGCTGAAGCTTCAAGCTATCAAGTCCACACCG 1167
Db 781 TCAATGTGACTGTGATGCCACTGTTGCTGAAGCTTCAACATATCAAGACCCACCA 840
QY 1168 CCTTATCCAAACATCTGTCTACACCAAAACCCCTCTGCTCCCTCATATGAGTCAAGTAA 1227
Db 841 CCTTATCCAAACATCTGTCTGACCAAAACCCCTCTGCTCCCTCATATGAGTCAAGTAA 900
QY 1228 AGCCCTGCAAGATGAACAGCCCTAGCTTTACCAAGAGATGATAGTGAAGAGTGGG 1287
Db 901 AGCCCTGCAAGAGAGGATCAGCAAGCTTGCCTCAAGAGAGTGAAGAGTAAAT 960
QY 1288 GACAGTGGTCACTCTGCGGATTAAGAAAACACATTAACACTTCACTTCACTGTT 1347
Db 961 GAAATGTTGATAGTGGGATTAAGAAAACACATTAACACTTCACTTCACTGTT 1020
QY 1348 CGGAAAAACAAGAGATGAAGAACGAAGAGAGTCTCGGATTTCAGAGTTACTCCCCACAG 1407

```

Db 1021 AGAAAAACAAGAGATGAAGGGAATCTCGTATTCAAAGTTATTCTCTCAA 1080
 QY 1408 GCCTTTAAAGTCTTTCAGAGACGACGCTAGAGACGCTCTCAAGTCTCATCAGCAGCT 1467
 Db 1081 GCATTTAAATCTTTATGAGCAACATGATGAGAAATGTACTCAATCTCATCAGCAGCT 1140
 QY 1468 CTGCATCGAAGACGACGCTAGAAAATGAATGATCGGGTGGGATTAATCTCAAGATGCC 1527
 Db 1141 CTACATCGTAAAAAACAATTAGAATGAATGATCGGGTGGGATTAATCTCAAGATGCC 1200
 QY 1528 CAGGATCAATAGAAAGATGCTTTCCGAGAAAGCTCTAACTATATTCGTCTTAAAGG 1587
 Db 1201 CAGATCAATAGAAAGATGCTTTCCGAGAAAGATCTAAATACATCCGCTCTTAAAGG 1260
 QY 1588 GCTAAATGGACAAAGCTATGTTTGAAGATGAAGATAGGATAGGAGGTTTGGT 1647
 Db 1261 GCTAAATGGACAAAGCTATGTTTGAAGATGAAGATAGGATAGGAGGTTTGGT 1320
 QY 1648 GAAGTCTGTCTAGCAAGAAAGTCGATCTAAAGCTTTGATGCAACAAAGACTCTTCGA 1707
 Db 1321 GAAGTCTGTCTAGCAAGAAAGTAGATCTAAGCTTTGATGCAACAAAGACTCTTCGA 1380
 QY 1708 AAGAAAGACGCTTCTGCTCCGAAATCAGGTGCTCAATGTAAGGAGGAGGATATCCTA 1767
 Db 1381 AAGAAAGATGCTTCTTCTCCGAAATCAAGTCCGCTCATCTTAAGGCTGAGAGATATCCTG 1440
 QY 1768 GCAGAACCCGACATGAGTGGTGGTCCGCTGTACTACTCTTCCAGGACAAAGGACAA 1827
 Db 1441 GCTGAAGCTGACAAATGAATGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1500
 QY 1828 TTGTACTTTGTGATGACTACATCTCTGGGGGGATATGATGACCTTATTAATAGAAATG 1887
 Db 1501 TTATACTTTGTAATGGACTACATCTCTGGGGGGATATGATGACCTTATTAATAGAAATG 1560
 QY 1888 GGCATCTTTCTGAAATCTGGCAGGATCTACATAGCAGAACTTACTGTGCAAGTTGAA 1947
 Db 1561 GGCATCTTTCTGAAATCTGGCAGGATCTACATAGCAGAACTTACTGTGCAAGTTGAA 1620
 QY 1948 AGTGTCTATAAATGGGTTTATTCATAGATATTAACCTGTATAACATTTTGAATGAC 2007
 Db 1621 AGTGTCTATAAATGGGTTTATTCATAGATATTAACCTGTATAACATTTTGAATGAC 1680
 QY 2008 CGTATGCGCCATATTAATGACTGACCTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2067
 Db 1681 CGTATGCGCCATATTAATGACTGACCTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1740
 QY 2068 GACTCCAAGTACTACAGAGTGGGATCACCCAGGCAAGATAGCATGATTTTCAATGAC 2127
 Db 1741 GATTTCTAAGTACTATCAGAGTGGTGGATCCACCGCAAGATAGCATGATTTTCAATGAC 1800
 QY 2128 GAATGGGAGATCTTCCAAATGTCGGTGGTGGGACAGACTGAAGCCACTGGAGCGGAGA 2187
 Db 1801 GAATGGGAGATCTTCCAAATGTCGGTGGTGGGACAGACTGAAGCCACTGGAGCGGAGA 1860
 QY 2188 GCTGTCTGCGCAGCAGCAGGATGCTAGCCCATCTCTGGTGGGACCTCCCAATATAT 2247
 Db 1861 GCTGTCTGCGCAGCAGCAGGATGCTAGCCCATCTCTGGTGGGACCTCCCAATATAT 1920
 QY 2248 GCACCTGAAGTCTACTGCGACAGATATACACAGCTGTGTGACTGGTGGAGTGTGTGT 2307
 Db 1921 GCACCTGAAGTGTGTGCTAGACAGATATACACAGTGTGTGACTGGTGGAGTGTGTGT 1980
 QY 2308 GTTATTTCTTTGTAATGTTGGTGGGACAACTCTCTTCTGGCACAACCCCATATAGAA 2367
 Db 1981 GTTATTTCTTTGTAATGTTGGTGGGACAACTCTCTTCTGGCACAACCCCATATAGAA 2040
 QY 2368 ACACAAATGAAGTTATCATCTGGCAAACTCTCTACACATCCCTCTCAAGTAACTG 2427
 Db 2041 ACACAAATGAAGTTATCAACTGGCAAACTCTCTACACATCCCTCTCAAGTAACTG 2100
 QY 2428 AGTCTCTGAAGCTCTCACTCATTTATCAAACTGTGCGAGGACGAGAACCCGCTCGGC 2487

Db 2101 AGTCTGAAGCTTCTGATCTTATTATTAAACTTTGCCGAGGACCCGAAGATCGGTAGGC 2160
 QY 2488 AGAARCGGTGCTGATGATGAATAAGGCTCATCTTTTAAAGACCATCGATTCTCTAGT 2547
 Db 2161 AGAATGCTGCTGATGAATAAGGCTCATCTTTTAAACAAATGATTCTCTCCAGT 2220
 QY 2548 GATCTGAGACAGCAGCTGCTTCATACATCCCTAAATCAGCATCCCAACAGATACATCA 2607
 Db 2221 GACCTGAGACAGCAGCTGCTTCATACATCTCTAAATCAGCATCCCAACAGATACATCA 2280
 QY 2608 AATTTCCGACCTTGTGATCTCTGATATAAATGTTGGAGGATGGCAGGAGGAGAAATATC 2667
 Db 2281 AATTTGATCTGTTGATCTCTGATATAAATGTTGGAGGATGGCAGGAGGAGAAATATC 2340
 QY 2668 ATGACACTCT-GAGCGGATGTTAAATGGAAGCACCCTCCGAGCAGCGCTTCTCTATGA 2726
 Db 2341 AATGACACTCTCAATGGGATGTTAAATGGAAGCACCCTCCGAGCAGCGCTTCTCTATGA 2400
 QY 2727 GTTCACTTTCCGAGGTTTGTGATGACAAATGCTACCCATA 2768
 Db 2401 AATTACCTTCCGAGGTTTGTGATGACAAATGCTACCCATA 2442

RESULT 7
 AL583963/c
 LOCUS
 DEFINITION
 Human DNA sequence from clone RP1-203A15 on chromosome 6, complete sequence.
 ACCESSION AL583963
 VERSION AL583963.11 GI:14547269
 KEYWORDS HTG.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 80146)
 Direct Submission
 Submitted (23-JUN-2001) Sanger Centre, Hinxton, Cambridgeshire, UK. E-mail enquiries: humquery@sanger.ac.uk
 CB101 ISA, UK. E-mail enquiries: humquery@sanger.ac.uk
 requests: clonerequest@sanger.ac.uk
 On Jun 25, 2001 this sequence version replaced gi:14133045.
 During sequence assembly data is compared from overlapping clones. Where differences are found these are annotated as variations together with a note of the overlapping clone name. Note that the variation annotation may not be found in the sequence submission corresponding to the overlapping clone, as we submit sequences with only a small overlap as described above.
 This sequence was finished as follows unless otherwise noted: all regions were either double-stranded or sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by at least one plasmid subclone or more than one M13 subclone; and the assembly was confirmed by restriction digest. The following abbreviations are used to associate primary accession numbers given in the feature table with their source databases: Em:, EMBL; Sw:, SWISSPROT; Tr:, TrEMBL; Wp:, WORMPEP; information on the WORMPEP database can be found at
 http://www.sanger.ac.uk/Projects/C_elegans/wormep/ This sequence was generated from part of bacterial clone contigs of human chromosome 6, constructed by the Sanger Centre Chromosome 6 Mapping Group. Further information can be found at
 http://www.sanger.ac.uk/HGP/Chr6
 RP1-203A15 is from the library RPCI-1 constructed by the group of Pieter de Jong. For further details see
 http://www.chori.org/bacpac/home.htm
 VECTOR: pCYPAC2
 IMPORTANT: This sequence is not the entire insert of clone RP1-203A15. It may be shorter because we sequence overlapping sections only once, except for a 100 base overlap.
 The true left end of clone RP1-317N9 is at 78147 in this sequence.
 The true right end of clone RP1-703H16 is at 2000 in this

FEATURES	source	sequence.	Location/Qualifiers
repeat_region			15966..16070
repeat_region			/note="FLAM_C repeat: matches 1..105 of consensus"
repeat_region			16076..16112
repeat_region			/note="Alu repeat: matches 79..160 of consensus"
repeat_region			16122..17029
repeat_region			/note="L1MA4A repeat: matches 5319..6300 of consensus"
repeat_region			17080..17261
repeat_region			/note="MER74A repeat: matches 1..171 of consensus"
repeat_region			17490..17788
repeat_region			/note="Alusg repeat: matches 1..296 of consensus"
repeat_region			17966..18263
repeat_region			/note="Alu repeat: matches 1..298 of consensus"
repeat_region			18502..18551
repeat_region			/note="25 copies 2 mer ta 100% conserved"
repeat_region			18552..18818
repeat_region			/note="Alusg repeat: matches 34..303 of consensus"
repeat_region			18929..19117
repeat_region			/note="MIR repeat: matches 41..241 of consensus"
repeat_region			19252..19557
repeat_region			/note="Alusg repeat: matches 1..306 of consensus"
repeat_region			19558..19843
repeat_region			/note="Alu repeat: matches 12..294 of consensus"
repeat_region			20971..21270
repeat_region			/note="Alusx repeat: matches 1..300 of consensus"
repeat_region			21302..21391
repeat_region			/note="FRAM/FAM repeat: matches 4..72 of consensus"
repeat_region			22015..22173
repeat_region			/note="MER5B repeat: matches 7..176 of consensus"
repeat_region			22419..23403
repeat_region			/note="Cpg island"
repeat_region			22531..22618
repeat_region			/note="44 copies 2 mer cc 64% conserved"
repeat_region			23378..23453
repeat_region			/note="MER5A repeat: matches 108..185 of consensus"
repeat_region			23686..23896
repeat_region			/note="Alusg/x repeat: matches 89..310 of consensus"
repeat_region			24365..24495
repeat_region			/note="FLAM_C repeat: matches 2..125 of consensus"
repeat_region			24659..24859
repeat_region			/note="MER3 repeat: matches 2..209 of consensus"
repeat_region			25072..25249
repeat_region			/note="Alusx repeat: matches 118..290 of consensus"
repeat_region			25256..25572
repeat_region			/note="Alusx repeat: matches 1..309 of consensus"
repeat_region			25575..25874
repeat_region			/note="Alusx repeat: matches 1..299 of consensus"
repeat_region			26374..26417
repeat_region			/note="22 copies 2 mer ta 100% conserved"
repeat_region			27253..27549
repeat_region			/note="Alusg repeat: matches 1..298 of consensus"
repeat_region			27551..27631
repeat_region			/note="LIME repeat: matches 5699..5778 of consensus"
repeat_region			27852..27915
repeat_region			/note="L2 repeat: matches 2203..2267 of consensus"
repeat_region			28020..28320
repeat_region			/note="Alusx repeat: matches 1..305 of consensus"
repeat_region			28891..29184
repeat_region			/note="Alusg repeat: matches 1..292 of consensus"
repeat_region			29260..29548
repeat_region			/note="Alusg repeat: matches 1..305 of consensus"
repeat_region			29567..29652
repeat_region			/note="Alu repeat: matches 202..287 of consensus"
repeat_region			29818..30056
repeat_region			/note="Alusg repeat: matches 1..293 of consensus"
repeat_region			30267..30570
repeat_region			/note="Alusg repeat: matches 1..302 of consensus"
repeat_region			30581..30762
repeat_region			/note="LIMEC repeat: matches 2353..2196 of consensus"
repeat_region			30827..31119
repeat_region			/note="Alu repeat: matches 1..312 of consensus"
repeat_region			31127..31259
repeat_region			/note="LIMEC repeat: matches 2112..2246 of consensus"
repeat_region			1298..1421
repeat_region			/note="L1MC5 repeat: matches 7790..7931 of consensus"
repeat_region			1298..1421
repeat_region			/note="L2 repeat: matches 2234..2359 of consensus"
repeat_region			2135..2886
repeat_region			/note="Cpg island"
repeat_region			2728..3050
repeat_region			/note="Alu repeat: matches 1..303 of consensus"
repeat_region			3174..3476
repeat_region			/note="Alu repeat: matches 2..303 of consensus"
repeat_region			3512..3956
repeat_region			/note="L2 repeat: matches 2264..2750 of consensus"
repeat_region			3975..4127
repeat_region			/note="Alusg/x repeat: matches 145..297 of consensus"
repeat_region			4132..4469
repeat_region			/note="Alusx repeat: matches 1..304 of consensus"
repeat_region			4470..4856
repeat_region			/note="HURPS-P3 repeat: matches 6797..7185 of consensus"
repeat_region			4850..6348
repeat_region			/note="HURPS-P3 repeat: matches 5120..5625 of consensus"
repeat_region			6349..6647
repeat_region			/note="Alusg repeat: matches 1..301 of consensus"
repeat_region			6648..6815
repeat_region			/note="HURPS-P3 repeat: matches 4954..5120 of consensus"
repeat_region			6840..6911
repeat_region			/note="HURPS-P3 repeat: matches 4691..4762 of consensus"
repeat_region			6912..6953
repeat_region			/note="LIPBa repeat: matches 742..701 of consensus"
repeat_region			6974..7347
repeat_region			/note="HURPS-P3 repeat: matches 4568..4942 of consensus"
repeat_region			7349..7384
repeat_region			/note="LIP repeat: matches 1790..1825 of consensus"
repeat_region			7382..7745
repeat_region			/note="LIPBa repeat: matches 1590..1958 of consensus"
repeat_region			7745..8418
repeat_region			/note="LIM2 repeat: matches 84..638 of consensus"
repeat_region			8458..8921
repeat_region			/note="LIM2 repeat: matches 677..198 of consensus"
repeat_region			9470..9751
repeat_region			/note="Alusx repeat: matches 12..293 of consensus"
repeat_region			9941..10043
repeat_region			/note="Alu repeat: matches 210..306 of consensus"
repeat_region			10059..10260
repeat_region			/note="Alusg/x repeat: matches 108..309 of consensus"
repeat_region			10295..10994
repeat_region			/note="LIPB1 repeat: matches 4004..4703 of consensus"
repeat_region			10995..11289
repeat_region			/note="Alusg repeat: matches 1..301 of consensus"
repeat_region			11290..12343
repeat_region			/note="LIPB1 repeat: matches 4703..5780 of consensus"
repeat_region			12344..12479
repeat_region			/note="Alusx repeat: matches 1..142 of consensus"
repeat_region			12480..12769
repeat_region			/note="Alu repeat: matches 1..290 of consensus"
repeat_region			12770..12957
repeat_region			/note="Alusx repeat: matches 142..309 of consensus"
repeat_region			12958..13331
repeat_region			/note="LIPB1 repeat: matches 5779..6152 of consensus"
repeat_region			13382..13962
repeat_region			/note="L1MA4A repeat: matches 3027..3603 of consensus"
repeat_region			13963..14255
repeat_region			/note="Alusx repeat: matches 1..306 of consensus"
repeat_region			14256..15963
repeat_region			/note="L1MA4A repeat: matches 3603..5331 of consensus"

```

repeat_region 31453..31645
/note="L1M1 repeat: matches 1407..1614 of consensus"
repeat_region 32228..32673
/note="L1M1 repeat: matches 100..541 of consensus"
repeat_region 32688..32820
/note="FLAM_C repeat: matches 1..132 of consensus"

Query Match 36.4%; Score 1168.8; DB 9; Length 80146;
Best Local Similarity 85.9%; Pred. No. 2.2e-271;
Matches 1309; Conservative 0; Mismatches 212; Indels 3; Gaps 1;

QY 1 GTCACAACTTCAATTAACCGAAAGAGTGAAGAGTCTTAAGAGTCTCTAGTTCT 60
|||||
Db 57326 GTCCAGCAATCAGTTAACCGAAAGAGTGAAGAGTCTTAAGAGTCTCTAGTTCT 57267

QY 61 CAGACAGCGGCCCTCTCTAGAGAGAAATGGTTTATCGTTCTGAAGCCCACTCA 120
|||||
Db 57266 CAGAGCATGGCCGACCTAGAGAGAAATGGTTTATCGTTCTGAAGCCCACTCA 57207

QY 121 CAGCGGATGTAGAGAGCTCTGTCTGGATCGGCAATTCAGCATTTGCTCAAGCTCAC 180
|||||
Db 57206 CAGACAGATGTAGAGAGCTTTGTCTGGATCTGGTATATCAGCATTTGTTCAAGCTCAC 57147

QY 181 CAAGCAATGAGCAGAGAGTGAACCCGCCACCCACCTCAAGTATAGGAGTGTACTCT 240
|||||
Db 57146 CTTAGCAACGAGCAGAGAGTGAACCCGCCACCCACCTCAAGTATAGGAGTGTACTCTCT 57087

QY 241 CAGCAGCTCCGAGAGCCAGACCCACCTCCCGAGGACCACTCCCTCCCGCTCA 300
|||||
Db 57086 CCACCACTCCAGAGGCCAGAGTCCCGCTCCAGAGGATACAACTCCACCTCCCGCTCA 57027

QY 301 TGGGAACCAAGCTCTCAGACAAAGCGTACTCTGGGAACATGAGTACGTAATCTCCCGA 360
|||||
Db 57026 TGGGAACCAAGCTCTCAGACAAAGCGTACTCTGGGAACATGAGTACGTAATCTCCCGA 56967

QY 361 ATCTCCCTGTTCCACTGGGGCGTGGCAGAGGGTACCTCCACCACTCTTACACT 420
|||||
Db 56966 ATCTCTCTCCACTGGGGCGTGGCAGAGGGTATCTCTCCACCACTCTTCAACACT 56907

QY 421 TCTCCCATGATCCCTACGAGGCTCAGAGGCCATTAAGTTCTGTCTCCAGTTGGTGA 480
|||||
Db 56906 TCTCCCATGATCCCTACGAGGCTCAGAGGCCATTAAGTTCTGTCTCCAGTTGGTGA 56847

QY 481 CAACCATCATCATGACAGTACTAGCAAAATTTAACTTTACACGAGGCGACCTGGAGTT 540
|||||
Db 56846 CAACCAATCATCATGACAGTACTAGCAAAATTTAACTTTACACGAGGCGACCTGGAGTT 56787

QY 541 CAGATGGTGGTGTAGTCTGATTTTATCGTGACCAAAA---TGTCCTCCACTGGTTCT 597
|||||
Db 56786 CAGATGGTGTAGTCTGATTTTATCGTGACCAAAAATGTTGTCCTCTGGTGGCACT 56727

QY 598 GTGACTCGGAGCCACCACTCCATATCTCTGACCCGCTAATGGACAAAGCCCTCT 657
|||||
Db 56726 GTGATCGGAGCCACCACTCCATATCTCTGACAGCAGCTAATGGACAAAGCCCTCT 56667

QY 658 GCTTTACAAACAGGGGCTTCTGTGCTCCACCACTATTCGCCCAATGGAAGCTTCTCAG 717
|||||
Db 56666 GCTTTACAAACAGGGGATCTGTGCTCTCTGCTCATATACAAATGGAAGTATTCCTCAG 56607

QY 718 TCGATGATGTGCCCACAGGACAGTCAATACATGAGCTTATATATATGTCCT 777
|||||
Db 56606 TCTATGATGTGCCCACAGGAAATAGTCAATACATGAGCTTATATATATGTCCT 76547

QY 778 GGACTGCAAAACAGCCTGGCCCGCAGTCTCTCTGCTCTGCGAGTCAATCCCAAGCGGT 837
|||||
Db 56546 GGACTGCAAAACAAATGGCCCTCAGTCTCTCTCTGCTCTGCGAGTCAATCCCGAGCAGT 56487

QY 838 GGGCATGAATTCCTACATGGCAACCTTAACATACCATGAGTCAATTTCTTAATAAC 897
|||||
Db 56486 GGGCATGAATTCCTACATGGCAACCTTAACATACCATGAGTCAATTTCTTAATAAC 56427

QY 898 CCATTAGGAAGTACAGCAAGTCACTCTGCTAAATCTCAGGCTTCTGCGCACTACAGTCACT 957
|||||

```

```

Db 56426 CCATTAGGAAATAGAGCAAGTCACTCTGTCTAATCTCAGCCTTCTGTACAAACACTCACT 56367
QY 958 GCATCACACCGCTCTCTTATTCACACAGCCGCTGAAAAGCATCGCGTCTGTGAACCCAGAG 1017
|||||
Db 56366 GCAATTCACACAGCTCTCTTATTCACACAGCTGTGAAAAGTATCGTGTATTAACCCAGAG 56307
QY 1018 CTGCAGACTGCTTTAGCCCCAACCCATCTCTTGGATGTCACAGCCAGTTCAGACTGTT 1077
|||||
Db 56306 CTACAGACTGCTTTAGCACCTACACACCTCTCTTGGATACACAGCAATTCACAACTGTT 56247
QY 1078 CAGCTTACCCCTTTTCTGAGGTACAGTCTCAAGTGTGCTGTATCCACCTCTGTGCT 1137
|||||
Db 56246 CAACCCAGTCTTTTCTGAGGAACTCTTCAATGTGACTGTGTATGCCACTGTGCT 56187
QY 1138 GAAGCTTCAAGCTATCAAGGTCCACACCGCTTATCCAAAACATCTGTACACCAAAAC 1197
|||||
Db 56186 GAAGCTTCAAACTATCAAGGACCAACCACTCTACCAAAACATCTGTGTACCAAAAC 56127
QY 1198 CCATCTGCTCCCTCCATATGAGTCAGTAAGTAAAGCCCTGCAAGATGAACAGCTTACCT 1257
|||||
Db 56126 CCATCTGCTCCCTCCATATGAGTCAGTAAGTAAAGCCCTGCAAGATGAACAGCTTACCT 56067
QY 1258 CCAAGGAGATGATAGTGAAGAGTGGGACAGTGTGCTGTGGGATAAAGAAAG 1317
|||||
Db 56066 CCAAGGAGATGATAGTGAAGAGTGGGACAGTGTGCTGTGGGATAAAGAAAG 56007
QY 1318 AAACGATTACAACTTACCTATCACTGTTCGGAAGAAAGAGATGAAGACGAAGA 1377
|||||
Db 56006 AAACGATTACAACTTACCTATCACTGTTCGGAAGAAAGAGATGAAGACGAAGA 55947
QY 1378 GAGTCTCGGATTCAGAGTACTCCACAGGCTTTAAGTTCTTCATGAGCAGCAGCTA 1437
|||||
Db 55946 GAATCTCGTATTCAAAAGTTATTCTCTCAAGCATTTAAATTTTATGAGCAACATGTA 55887
QY 1438 GAGACGCTCTGAAGTCTCATCAGCAGCTGTCTGCGAAGAGCAGCTGAGAAATGAA 1497
|||||
Db 55886 GAAATGTACTCAATCTCATCAGCAGCTGTCTGCGAAGAGCAGCTGAGAAATGAA 55827
QY 1498 ATGATCGGGTGGATTTATCTCAA 1521
Db 55826 ATGATCGGGTAAACCTTTTAA 55803

RESULT 8
E24614 1374 bp DNA linear PAT 18-JUN-2001
LOCUS
DEFINITION
E24614 warts protein, polynucleotide encoding the protein, antisense
E24614 polynucleotide thereof and antibody recognizing the protein.
ACCESSION
E24614 1 GI:13024647
VERSION
E24614 1
KEYWORDS
E24614 1 JP 1999089580-A/2.
SOURCE
E24614 1 unidentified.
ORGANISM
E24614 1 unclassified.
REFERENCE
1 (bases 1 to 1374)
AUTHORS
Toshihiko, K., Shinichi, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
TITLE
warts protein, polynucleotide encoding the protein, antisense
JOURNAL
polynucleotide thereof and antibody recognizing the protein
SUMITOMO ELECTRIC IND LTD
COMMENT
OS Unidentified
PN JP 1999089580-A/2
PD 06-APR-1999
PF 24-SEP-1997 JP 1997258689
PR
PI TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
YASUYUKI NISHIYAMA,
PI HIDEYUKI SAYA
PC C12N15/09, C07K14/435, C07K16/18, C12Q1/68, G01N33/53, PC
G01N33/532, C12P21/02,
PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,
(C12N15/00, PC, C12R1:91)
CC Strandedness: Double;

```

D	B	841		CCCAATTATATTCACCTGAAGTGTGCTACTAGAACAGGATACACACAGTTGTGTGATTGG	900
Q	Y	2296		TGGAGTGCTTGGTGTTATCTTTTGTGAATGTTGGTGGGACAACCTCCTTTCTTGGGCACAA	2355
D	B	901		TGGAGTGCTTGGTGTTATCTTTTGTGAATGTTGGTGGGACAACCTCCTTTCTTGGGCACAA	960
Q	Y	2356		ACCCCATTAGAAAACACAAAATGAAGTTTATCATCTGGCAAACCTTCTCTACACATCCCTCCT	2415
D	B	961		ACACCATTAGAAAACACAAAATGAAGTTTATCAACTGGCAAACCTCTCTACATTCACCA	1020
Q	Y	2416		CACGCTAAGCTGAGTCTCGTAAGCTCTGACCTCATTTATCAAACTGTCGAGGACACGAA	2475
D	B	1021		CAAGCTTAACTCAGTCTGAGCTTCTGATCTTATTTATAAATTTGCCGAGAGACCCGAA	1080
Q	Y	2476		GAGCGCTCGGCAAGACGCTGCTGATGAGATAAAGCTCATCCATTTTTTAAGACCATC	2535
D	B	1081		GATCGCTTAGCAAGAATGCTGCTGATGAATAAAGCTCATCCATTTTTTAACAACTT	1140
Q	Y	2536		GATTTCTCTAGTCNGAGACAGCAGTCTGCTTCATATACATCCCTAAAAATCACGCAPTCA	2595
D	B	1141		GATTTCTCCAGTACCTTGACAGACAGCAGTCTGCTTCATACATCTCTAAATACACACCCA	1200
Q	Y	2596		ACAGATACATCCCAATTTTCGACCTGTGTGATCCTGATATAAATTTGGAGCGATGGCAGGAG	2655
D	B	1201		ACAGATACATCAAAATTTTGATCTGTTGATCCTGATATAAATTTGGAGCTGATGATAAGAG	1260
Q	Y	2656		GAGGAAATATCAGTGACACTCT - GAGCGGATGATATAAATTTGGGAAGCACCCCGAGCA	2714
D	B	1261		GAGGAAATGTAATGACACTCTCATGGGATGATATAAATTTGGGAAGCATCTGTAACA	1320
Q	Y	2715		CGCTTTCTATGAGTTCACCTTTTCGGAGGTTTTTGATGACAAATGGCTACCCATA	2768
D	B	1321		TGCATTCTATGAATTTACCTTCGGAAGTTTTTTTGATGACAAATGGCTACCCATA	1374
RESULT 9					
AC102712/c					
LOCUS					
DEFINITION					
ACCESSION					AC102712
VERSION					164490 bp DNA linear HTG. 21-AUG-2002
KEYWORDS					Mus musculus clone RP24-25894, WORKING DRAFT SEQUENCE, 17 unordered pieces.
SOURCE					AC102712.2 GI:22381704
ORGANISM					HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP. house mouse. Mus musculus Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 164490) Birren,B., Nusbaum,C. and Lander,E. Mus musculus, clone RP24-25894 Unpublished 2 (bases 1 to 164490) Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhgalter,B., Brown,A., Camarata,J., Campiolino,A., Chang,J., Chazarro,B., Choepel,Y., Colangelo,M., Collins,S., Collamore,A., Cook,A., Cooke,P., DeAtellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., Fitzhugh,W., Gage,D., Galagan,J., Gardyna,S., Ginde,S., Gord,S., Guyette,M., Graham,L., Grand-Pierre,N., Hagos,B., Hearford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kamat,A., Karatas,A., Kells,C., LaRoque,K., Lamazares,R., Landers,T., Lehoczyk,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., MCPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Miengva,V., Murphy,T., Naylor,J., Nguyen,C., Norbu,C., Norman,C.H., O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Raymond,C., Retta,R., Riback,M., Riley,R., Rise,C., Rogov,P., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupback,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strauss,N., Subramanian,A., Talamas,J., Testaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Trigilio,J., Vassiliev,H., Viet,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G.

TITLE
JOURNAL
REFERENCE
AUTHORS

zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
3 (bases 1 to 164490)
Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S.,
Barna,N., Bastien,V., Bloom,T., Boguski,M., Bouckgeater,B.,
Cammarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A.,
Cook,A., Cooke,P., DeArillano,K., Dewar,K., Diaz,J.S., Dodge,S.,
Faro,S., Ferreira,P., Fitzgerald,M., Gage,D., Galagan,J.,
Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hagos,B.,
Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A.,
Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K.,
Liu,G., Maclean,C., Macdonald,P., Major,J., Matthews,C.,
McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V.,
Murphy,T., Navlor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H.,
O'Connor,T., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K.,
PhunKhang,P., Pierre,N., Raymond,C., Retta,R., Rise,C., Rogov,P.,
Roman,J., Roy,A., Schauer,S., Schupback,R., Seaman,S., Severy,P.,
Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J.,
Tessfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H.,
Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J.,
Zembek,L., Zimmer,A. and Zody,M.
Direct Submission
Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome
Research, 320 Charles Street, Cambridge, MA 02141, USA
On Aug 21, 2002 this sequence version replaced gi:17060822.
All repeats were identified using RepeatMasker:
Smit, A.F.A. & Green, P. (1996-1997)
<http://ftp.genome.washington.edu/RM/RepeatMasker.html>
----- Genome Center
Center: Whitehead Institute/ MIT Center for Genome Research
Center code: WIBR
Web site: <http://www-seq.wi.mit.edu>
Contact: sequence_submissions@genome.wi.mit.edu
----- Project Information
Center project name: L19333
Center clone name: 258.P.4
----- Summary Statistics
Sequencing vector: Plasmid; n/a; 100% of reads
Chemistry: Dye-terminator Big Dye; 100% of reads
Assembly program: Phrap; version 0.960731
Consensus quality: 158929 bases at least Q40
Consensus quality: 161241 bases at least Q30
Consensus quality: 162174 bases at least Q20
Insert size: 176000; agarose-fp
Quality coverage: 7.2 in Q20 bases; agarose-fp
Quality coverage: 7.8 in Q20 bases; sum-of-contigs

* NOTE: This is a 'working draft' sequence. It currently
* consists of 17 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.
*
* 1 1303: contig of 1303 bp in length
* 1304 1403: gap of 100 bp
* 1404 2163: contig of 760 bp in length
* 2164 2263: gap of 100 bp
* 2264 3531: contig of 1268 bp in length
* 3532 3631: gap of 100 bp
* 3632 4716: contig of 1085 bp in length
* 4717 4816: gap of 100 bp
* 4817 8090: contig of 3274 bp in length
* 8091 8190: gap of 100 bp
* 8191 11116: contig of 2926 bp in length
* 11117 11216: gap of 100 bp
* 11217 41036: contig of 29820 bp in length
* 41037 41136: gap of 100 bp
* 41137 45062: contig of 3926 bp in length

TITLE
JOURNAL
COMMENT

* 45063 45162: gap of 100 bp
* 45163 48629: contig of 3467 bp in length
* 48630 48729: gap of 100 bp
* 48730 57775: contig of 9046 bp in length
* 57776 57875: gap of 100 bp
* 57876 70851: contig of 12976 bp in length
* 70852 70951: gap of 100 bp
* 70952 85997: contig of 15046 bp in length
* 85998 86097: gap of 100 bp
* 86098 104897: contig of 18800 bp in length
* 104898 104997: gap of 100 bp
* 104998 120197: contig of 15200 bp in length
* 120198 120297: gap of 100 bp
* 120298 136805: contig of 16508 bp in length
* 136806 136905: gap of 100 bp
* 136906 160104: contig of 23199 bp in length
* 160105 160204: gap of 100 bp
* 160205 164490: contig of 4286 bp in length.
Location/Qualifiers
1. 164490
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone_lib="RPC1-24 Male Mouse BAC"
1. 1303
/note="assembly_fragment"
1404. 2163
/note="assembly_fragment"
2264. 3531
/note="assembly_fragment"
3632. 4716
/note="assembly_fragment"
4817. 8090
/note="assembly_fragment"
8191. 11116
/note="assembly_fragment"
11217. 41036
/note="assembly_fragment"
41137. 45062
/note="assembly_fragment"
45163. 48629
/note="assembly_fragment"
48730. 57775
/note="assembly_fragment"
57876. 70851
/note="assembly_fragment"
70952. 85997
/note="assembly_fragment"
86098. 104897
/note="assembly_fragment"
104998. 120197
/note="assembly_fragment"
120298. 136805
/note="assembly_fragment"
136906. 160104
/note="assembly_fragment"
160205. 164490
/note="assembly_fragment"
clone_end:77
vector_side:right
BASE COUNT 47330 a 35286 c 34742 g 45505 t 1627 others
ORIGIN
Query Match 33.6%; Score 1079.4; DB 2; Length 164490;
Best Local Similarity 99.8%; Pred. No. 9.6e-250;
Matches 1091; Conservative 0; Mismatches 1; Indels 1; Gaps 1;
QY 1 GTGCAACATTCAATTAAACCGAAACAAAGCTGGAAGCTTCTAAAGAGTCTTAGTTCCT 60
|||||
Db 5913 GTGCAACATTCAATTAAACCGAAACAAAGCTGGAAGCTTCTAAAGAGTCTTAGTTCCT 5854
QY 61 CAGAGACACGCCCTCTCTAGGAGAAATGTGGTTTATCGTTCTGAAAGCCCCCAACTCA 120
|||||

Db 5853 CAGAGACAGCGCCCATCTCTAGGAGAAAATGTGTTTATCTGTCTGTAAGAGCCCAACTCA 5794
 Qy 121 CAGCGGATGTAGGAAGACCTCTGTCTGATCCGCAATTCAGAGCTTGTCTCAAGCTCAC 180
 Db 5793 CAGCGGATGTAGGAAGACCTCTGTCTGATCCGCAATTCAGAGCTTGTCTCAAGCTCAC 5734
 Qy 181 CCAAGCAATGTAGGAGAGTGAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 240
 Db 5733 CCAAGCAATGTAGGAGAGTGAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCAACCCCA 5674
 Qy 241 CCACCACTCCAGAGGAGCCAGACCCCACTCCAGAGGAGCCCACTCCAGAGGAGCCCACTCCAGAGG 300
 Db 5673 CCACCACTCCAGAGGAGCCAGACCCCACTCCAGAGGAGCCCACTCCAGAGGAGCCCACTCCAGAGG 5614
 Qy 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGAACATGAGTACGTAATCTCCCGA 360
 Db 5613 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGAACATGAGTACGTAATCTCCCGA 5554
 Qy 361 ATCTCCCTGTTCACCTGGGCGTGGCAGGAGGGTACCCCTCCACCACTCTTACCCT 420
 Db 5553 ATCTCCCTGTTCACCTGGGCGTGGCAGGAGGGTACCCCTCCACCACTCTTACCCT 5494
 Qy 421 TCTCCCATGAATCCCTCTAGCAGGCTCAGAGGCGCATAGTCTGTCCAGTTCGTAGA 480
 Db 5493 TCTCCCATGAATCCCTCTAGCAGGCTCAGAGGCGCATAGTCTGTCCAGTTCGTAGA 5434
 Qy 481 CAACCCATCATCATGAGTACTAGCAAAATTTAACTTTACACCAAGGCGACCTCGAGTT 540
 Db 5433 CAACCCATCATCATGAGTACTAGCAAAATTTAACTTTACACCAAGGCGACCTCGAGTT 5374
 Qy 541 CAGAAATGTGTGTGCTGATTTATTCGTGACCAAAATGTCCTCCACTGTTCTGTG 600
 Db 5373 CAGAAATGTGTGTGCTGATTTATTCGTGACCAAAATGTCCTCCACTGTTCTGTG 5314
 Qy 601 ACTCGGACGACCACTCCATATCTCTGACCCCAAGTAAAGCAAGCCCTCTGCT 660
 Db 5313 ACTCGGACGACCACTCCATATCTCTGACCCCAAGTAAAGCAAGCCCTCTGCT 5254
 Qy 661 TTACAAACAGGGGCTTCTGCTGCTCCTACCATCATTCGCAATGGAAGCTTCTCTCAGTCG 720
 Db 5253 TTACAAACAGGGGCTTCTGCTGCTCCTACCATCATTCGCAATGGAAGCTTCTCTCAGTCG 5194
 Qy 721 ATGATGTGTGCGCAACAGGAACAGTCAATACATGAGCTTTATATTAATGTCCTGGA 780
 Db 5193 ATGATGTGTGCGCAACAGGAACAGTCAATACATGAGCTTTATATTAATGTCCTGGA 5134
 Qy 781 CTGCAACAGGCTGCGCCCAAGTCTCTGCTGCTGCGCAGTCAATCCCAAGCGGTGG 840
 Db 5133 CTGCAACAGGCTGCGCCCAAGTCTCTGCTGCTGCGCAGTCAATCCCAAGCGGTGG 5074
 Qy 841 CATGAATTCCTACATGCAACCTAACATACAGTGAAGTCAATTCCTTTAATACCCA 900
 Db 5073 CATGAATTCCTACATGCAACCTAACATACAGTGAAGTCAATTCCTTTAATACCCA 5014
 Qy 901 TTAGGAAGTAGGCAAGTCACTCTGCTAATTCAGCCCTTCTGCCACTACAGTCACTGCC 960
 Db 5013 TTAGGAAGTAGGCAAGTCACTCTGCTAATTCAGCCCTTCTGCCACTACAGTCACTGCC 4954
 Qy 961 ATCAGACCGCTCTCTATTCAGACGCGGTGAAGACATGCGGCTCTGAAACCAAGCTG 1020
 Db 4953 ATCAGACCGCTCTCTATTCAGACGCGGTGAAGACATGCGGCTCTGAAACCAAGCTG 4894
 Qy 1021 CAGACTGCTTTAGCCCAACCCATCTCTTGTGATGCCACAGCCAGTTCAGACTGTTACG 1080
 Db 4893 CAGACTGCTTTAGCCCAACCCATCTCTTGTGATGCCACAGCCAGTTCAGACTGTTACG 4835
 Qy 1081 CCTACCCCTTTT 1093
 Db 4834 CCTACCCCTTTT 4822

RESULT 10
 AC102712

LOCUS AC102712 164490 bp DNA linear HTG 21-AUG-2002
 DEFINITION Mus musculus clone RP24-258P4, WORKING DRAFT SEQUENCE, 17 unordered pieces.
 ACCESSION AC102712
 VERSION AC102712.2 GI:22381704
 KEYWORDS HTG; HTGS_PHASE1; HTGS_DRAFT; HTGS_FULLTOP.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 Birren,B., Nusbaum,C. and Lander,E.
 1 (bases 1 to 164490)
 Mus musculus, clone RP24-258P4
 Unpublished
 2 (bases 1 to 164490)
 Birren,B., Linton,L., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Boguslavsky,L., Boukhalter,B., Brown,A., Camarata,J., Campiano,A., Chang,J., Chazaro,B., Choepel,Y., Colangelo,M., Collins,S., Collymore,A., Cook,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzHugh,W., Gage,D., Galagan,J., Gardyna,S., Ginder,S., Gord,S., Goyette,M., Graham,L., Grand-Pierre,N., Hago,S., Heaford,A., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., LaRocque,K., Lamazares,R., Landers,T., Lehoczy,J., Levine,R., Liu,G., MacLean,C., Macdonald,P., Major,J., Marquis,N., Matthews,C., McCarthy,M., McEwan,P., McKernan,K., McPheeters,R., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Norman,C., Norbu,C., O'Connor,P., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Pollara,V., Roman,J., Rosetti,M., Roy,A., Santos,R., Schauer,S., Schupack,R., Seaman,S., Severy,P., Spencer,B., Stange-Thomann,N., Stojanovic,N., Strausman,A., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Travis,N., Triggillo,J., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Ye,W.J., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (23-NOV-2001) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 3 (bases 1 to 164490)
 Birren,B., Nusbaum,C., Lander,E., Ali,A., Allen,N., Anderson,S., Barna,N., Bastien,V., Bloom,T., Boguslavsky,L., Boukhalter,B., Camarata,J., Chang,J., Chazaro,B., Choepel,Y., Collymore,A., Cooke,A., Cooke,P., DeArellano,K., Dewar,K., Diaz,J.S., Dodge,S., Faro,S., Ferreira,P., FitzGerald,M., Gage,D., Galagan,J., Gardyna,S., Gord,S., Graham,L., Grand-Pierre,N., Hago,S., Horton,L., Hulme,W., Iliev,I., Johnson,R., Jones,C., Kanat,A., Karatas,A., Kells,C., Landers,T., Levine,R., Lindblad-Toh,K., Liu,G., MacLean,C., Macdonald,P., Major,J., Matthews,C., McCarthy,M., Meldrum,J., Meneus,L., Mihova,T., Mlenga,V., Murphy,T., Naylor,J., Nguyen,C., Nicol,R., Norbu,C., Norman,C.H., O'Connor,P., O'Donnell,P., O'Neill,D., Oliver,J., Peterson,K., Phunkhang,P., Pierre,N., Raymond,C., Retta,R., Risse,C., Rogov,P., Roman,J., Roy,A., Schauer,S., Schupack,R., Seaman,S., Severy,P., Smith,C., Spencer,B., Stange-Thomann,N., Stojanovic,N., Talamas,J., Tesfaye,S., Theodore,J., Topham,K., Travers,M., Vassiliev,H., Viel,R., Vo,A., Wilson,B., Wu,X., Wyman,D., Young,G., Zainoun,J., Zembek,L., Zimmer,A. and Zody,M.
 Direct Submission
 Submitted (21-AUG-2002) Whitehead Institute/MIT Center for Genome Research, 320 Charles Street, Cambridge, MA 02141, USA
 On Aug 21, 2002 this sequence version replaced gi:17060822.
 All repeats were identified using RepeatMasker:
 Smit, A.F.A. & Green, P. (1996-1997)
 http://ftp.genome.washington.edu/RM/RepeatMasker.html
 ----- Genome Center
 Center: Whitehead Institute/ MIT Center for Genome Research
 Center code: WIBR
 Web site: http://www-seq.wi.mit.edu
 Contact: sequence_submissions@genome.wi.mit.edu
 ----- Project Information
 Center project name: L19333

```

/note="assembly_fragment"
41137. .45062
/note="assembly_fragment"
45163. .48629
/note="assembly_fragment"
48730. .57775
/note="assembly_fragment"
57876. .70851
/note="assembly_fragment"
70952. .85997
/note="assembly_fragment"
86098. .104897
/note="assembly_fragment"
104998. .120197
/note="assembly_fragment"
120298. .136805
/note="assembly_fragment"
136906. .160104
/note="assembly_fragment"
160205. .164490
/note="assembly_fragment
clone_end:77
clone_side:right"

```

BASE COUNT	47330 a	35286 c	34742 g	45505 t	1627 others
ORIGIN	vector_side:right				
Query Match	27.2%; Score 872.4; DB 2; Length 164490;				
Best Local Similarity	99.8%; Pred. No. 1.1e-199;				
Matches 884; Conservative	0; Mismatches 1; Indels 1; Gaps 1;				
QY 624	TCCCTGTGACCCAGCTAATGACAAAGCCCTCTGCTTTTACAAACAGGGGCTCTGTGTCG	683			
Dbb 136969	TGCTGTGACCCAGCTAATGACAAAGCCCTCTGCTTTTACAAACAGGGGCTCTGTGTCG	137028			
QY 684	TCCACCATCATTCGCCAATGAAACGTTCTCTAGTGCATGATGGTGCCTCCACAGGAACAG	743			
Dbb 137029	TCCACCATCATTCGCCAATGAAACGTTCTCTAGTGCATGATGGTGCCTCCACAGGAACAG	137088			
QY 744	TCATAACATGGAGCTTTATATATATATGCTTCCCTGGACTGCAACAGCCTGGCCCCAGTTC	803			
Dbb 137089	TCATAACATGGAGCTTTATATATATATGCTTCCCTGGACTGCAACAGCCTGGCCCCAGTTC	137148			
QY 804	GTCTTCGTCTCTCGCGCAGTCAATCCCAAGCGGTGGGCATGAAATTCCTACATGGCAACC	863			
Dbb 137149	GTCTTCGTCTCTCGCGCAGTCAATCCCAAGCGGTGGGCATGAAATTCCTACATGGCAACC	137208			
QY 864	TAACATACCAAGTGAAGTCAAAATCTTTTAAACCCATTAGGAAGTAGACCAAGTCACATC	923			
Dbb 137209	TAACATACCAAGTGAAGTCAAAATCTTTTAAACCCATTAGGAAGTAGACCAAGTCACATC	137268			
QY 924	TGCTAATCTTCAGCCCTTCGCGCACTACAGTCTGCCATCACACCCCTCCTATTCAACA	983			
Dbb 137269	TGCTAATCTTCAGCCCTTCGCGCACTACAGTCTGCCATCACACCCCTCCTATTCAACA	137328			
QY 984	GCCCGTGAAGAAGCATCGCGTCTGNAACAGAGCTGCAGACTGCTTTAGCCCCAACCCA	1043			
Dbb 137329	GCCCGTGAAGAAGCATCGCGTCTGNAACAGAGCTGCAGACTGCTTTAGCCCCAACCCA	137388			
QY 1044	TCCTCTCTGGATGCCACAGCCAGTCTCAGACTGTTCAGCCTACCCCTTTTCTGAGGGTAC	1103			
Dbb 137389	TCCCTTCCTGGATGCCACAGCCAGTCTCAGACTGTTCAGCCTACCCCTTTTCTGAGGGTAC	137448			
QY 1104	AGCTTCAAGTGTGCTGTCAATCCCACTGTTGCTGAAGCTCCAAGCTATCAAGGTCCACC	1163			
Dbb 137449	AGCTTCAAGTGTGCTGTCAATCCCACTGTTGCTGAAGCTCCAAGCTATCAAGGTCCACC	137508			
QY 1164	ACCGCCTTATCCAAACATCTGCTTACACCAAAAC -GCATCTGTCCCTCCATGTAGTCTAG	1222			
Dbb 137509	ACCGCCTTATCCAAACATCTGCTTACACCAAAAC -GCATCTGTCCCTCCATGTAGTCTAG	137568			
QY 1223	TAAGTAAGCCCTGCAAGATGAACAGCCTAGCTTACCCAGGGAAGATGATAGTGAGAAGA	1282			
Dbb 137569	TAAGTAAGCCCTGCAAGATGAACAGCCTAGCTTACCCAGGGAAGATGATAGTGAGAAGA	137628			

FEATURES	source	Location/Qualifiers
	1.	.164490
		/organism="Mus musculus"
		/db_xref="taxon:10090"
		/clone="RP24-258P4"
		/clone_lib="RPCL-24 Male Mouse BAC"
	1.	.1303
misc_feature		/note="assembly_fragment"
	1404.	.2183
misc_feature		/note="assembly_fragment"
	2264.	.3531
misc_feature		/note="assembly_fragment"
	3632.	.4716
misc_feature		/note="assembly_fragment"
	4817.	.8090
misc_feature		/note="assembly_fragment"
	8191.	.11116
misc_feature		/note="assembly_fragment"
	11217.	.41036
misc_feature		


```
QY 1283 GTGCGACAGTGGTGAAGTCTGGGGATAAGAAAAAGAAAGAGATTACAACTTCCACCTATCA 1342
|||||
Db 137629 GTGCGACAGTGGTGAAGTCTGGGGATAAGAAAAAGAAAGAGATTACAACTTCCACCTATCA 137688

QY 1343 CTGTTCCGAAACAAAGAAAGATGAAGAACGAGAGAGTCTCGGATTCAGAGTTACTGCC 1402
|||||
Db 137689 CTGTTCCGAAACAAAGAAAGATGAAGAACGAGAGAGTCTCGGATTCAGAGTTACTGCC 137748

QY 1403 CACAGGCTTTAAAGTCTTCATGGAGCAGCAGTACAGAGACCTCTGAAGTCTCATCAGC 1462
|||||
Db 137749 CACAGGCTTTAAAGTCTTCATGGAGCAGCAGTACAGAGACCTCTGAAGTCTCATCAGC 137808

QY 1463 ACAGCTCGCATCGGAAGAGCAGCTAGAAAAATGAATGATGCGGGT 1508
|||||
Db 137809 ACAGCTCGCATCGGAAGAGCAGCTAGAAAAATGAATGATGCGGGT 137854

RESULT 11
AR201457
LOCUS AR201457 3155 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 7 from patent US 6359193.
ACCESSION AR201457
VERSION AR201457.1 GI:20252345
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE
1 (bases 1 to 3155)
AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes.
JOURNAL Patent: US 6359193-A 7 19-MAR-2002;
FEATURES
Location/Qualifiers
source 1..3155
BASE COUNT 751 a 924 c 894 g 586 t
ORIGIN
Query Match 25.3%; Score 812.6; DB 6; Length 3155;
Best Local Similarity 65.1%; Pred. No. 2.5e-185;
Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

QY 919 CACTCTGTCTAATTCACAGCTTCTGCCACTACAGTCACTGACCTACACACCCGCTCCTATT 978
|||||
Db 985 CAACCTGAGCCCTACTGCCGCCCCCAACAGCGTCAACCGCTGACGCCGCCACACATC 1044

QY 979 CAACAGCCCGTGAAGAGATGGCGTCTCTGAACAGAGCTGACAGCTGCTTTAGCCCA 1038
|||||
Db 1045 CTTACCCCTGTGAAGAGCGTGGTGTGTGCGGCCGAGCCCGACAGACGCGGTGGGGCCC 1104

QY 1039 ACCCATCTCTTGGATGCCACAGCAGTTCAGAGCTGTTACGCTTACCCCTACCCCTTTTCTGAG 1098
|||||
Db 1105 TGCACCCCGCTGGTGGCTGGCCCAACAGCACCTGCCACTGAGAGCGCTGGAGACGAG. 1164

QY 1099 GGTACAGCTTCAAGTGTGCTGTATCCACCTGTGCTGAAGCTTCAAGCTATCAAGT. 1158
|||||
Db 1165 GAGGCGACGCGAGGCCACACCCGCTGGATGTGAGTATGGCGCTCCGAGCGAGTGC 1224

QY 1159 CCACACCCGCTTATCCAAACATCTGTACACCAAAACCCATCTGTCCTCCATATGAG 1218
|||||
Db 1225 CCACCGCTCCGCTATCCAAAGACATCTGCTGTGCCAGTAAGTCTGAGCAGTACAGCGTG 1284

QY 1219 TCAGTAAGTAAGCCCTGCAAGATGAACAGCTACCTTACCCCAAGAGATGATAGTGA 1278
|||||
Db 1285 GACCTGGACAGCTGTGCACCAAGTGTGACGAGAGTCTGCGAGGGGCGACTGATCAGAC 1344

QY 1279 AAGAGTGGCGACAGT-----GGTGACTCTGGGGATAAAGAAAGAAACAG 1323
|||||
Db 1345 GCGAGTGACAGAGCCCAAGAGTGGAGGAGGACAAAGCTGGCAGAGACAAAGACAG 1404

QY 1324 ATTACAACCTTACCTATCAGTCTTCGGAACAAAGAGATGAAGAGACGAGAGTCT 1383
|||||
Db 1405 ATTACAGCTTCCCGGCTGCTGTCCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
```

```
QY 2464 CGAGGACCAAGACCGCCCTCGCAGACAGACGGTGTGATGAGATAAAGGCTCATCCATT 2523
Db 2545 TGGCGGCTGACTTGGCCCTGGCGAGGATGGGCGATGACCTCAAGGCACACCGCTTC 2604
QY 2524 TTTAAGACCATCATTTCTTCTAGTGTCTGAGACAGAGTCTCTCATACATCCCTAAA 2583
Db 2605 TTCAACACCATCATTTCTTCTAGTGTCTGAGACAGAGTCTCTCATACATCCCTAAA 2664
QY 2584 ATCAGCATCCCAAGATATATATATATATATATATATATATATATATATATATATAT 2643
Db 2665 ATCAGCATCCCAAGATATATATATATATATATATATATATATATATATATATATAT 2724
QY 2644 GATGAGCAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2703
Db 2725 GAGGCGAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2781
QY 2704 CACCCGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2763
Db 2782 CATCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2841
QY 2764 CCATATAATTATCAAGACCTATTGAGTATGATATATATATATATATATATATATATAT 2802
Db 2842 CCCTCCGGTGGCCGAGGAGCCTCAGAGCCGCGAGAGT 2880

RESULT 12
LOCUS AB023958 3460 bp mRNA linear ROD 09-MAR-2000
DEFINITION Mus musculus MmLATS2 mRNA for 'warts/lats'-like kinase, complete cds.
ACCESSION AB023958
VERSION MmLATS2.1 GI:7212785
KEYWORDS warts/lats-like kinase.
SOURCE Mus musculus
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (sites)
Yabuta, N., Fujii, T., Copeland, N.G., Gilbert, D.J., Jenkins, N.A.,
Nishiguchi, H., Endo, Y., Toji, S., Tanaka, H., Nishimune, Y. and
Nojima, H.
Structure, expression, and chromosome mapping of LATS2, a mammalian
homologue of the Drosophila tumor suppressor gene lats/warts
Genomics 63 (2), 263-270 (2000)
20139436
2 (bases 1 to 3460)
Nojima, H. and Fujii, T.
Direct Submission
Submitted (22-FEB-1999) Hiroshi Nojima, Research Institute for
Microbial Diseases, Osaka University, Department of Molecular
Genetics, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
(E-mail:hnojima@biken.osaka-u.ac.jp, Tel:81-6-6875-3980,
Fax:81-6-6875-5192)
Sequence updated (06-Jan-2000).
FEATURES
source
1..3460
/organism="Mus musculus"
/db_xref="taxon:10090"
1..3460
/gene="MmLATS2"
134..3262
/contig="MmLATS2"
/notes="WALK"
/codon_start=1
/product="warts/lats-like kinase"
/protein_id="BAA92380.1"
/db_xref="GI:7212786"
/translation="MRPKTFPATYSCNSRQRLQETREGKOPSKASTOGLVGNPSD
TSLDAKVGSKASROQMRKPKFGYONALREIYSLPFPANESCTSAAEVNRQM
LQELVNAVGDQEMAGRALQPSYRLEGEALYISKMGYLDPRNEQLVRVIKOTSPGK
LAPTVPYTRPSFEGTEALPSYRLEGEALYISKMGYLDPRNEQLVRVIKOTSPGK
QAHOHPKGYSTAVEPSTAFHFGTTHYGRHLLSEOPGYGVORSSFNKPTPPDAYSSMA
KAQGGPPASLTIPAHAGLYTASHHKPAATPGAHPLHVLGTRGPTFTGESSAAVLAP
SRNSLNADLYELGISTVPWSAAPLARDLSLQKQLEASRPHVAFRAGPSRTNSFNPPQ
```

polyA_site

/gene="MmLATS2"

BASE COUNT 861 a 992 c 963 g 644 t
ORIGIN

Query Match 25.1%; Score 806.2; DB 10; Length 3460;

Best Local Similarity 64.9%; Pred. No. 8.8e-184; Indels 18; Gaps 2;
Matches 1233; Conservative 0; Mismatches 648;

```
QY 919 CACTCTGCTAAATTTCTCAGCCTTCTGCCACTACAGTCTCAGTCCATCACACCCGCTCCTATT 978
Db 1304 CAACCTGAGCCCTCACTGCCGCCGCCACACAGCTCACCCGCTGACGCGCACATC 1363
QY 979 CAACAGCCGCTGAAAGCATGCGCGTCTGAAACACAGCTGCAGACTGCTTTAGCCCCA 1038
Db 1364 CTTACACCTGTGAAGACGCTGCGTGTGCGGCCGAGCCGACAGCCGCTGGGGCCC 1423
QY 1039 ACCATCTCTTGTGATGCCACAGCCAGTTCAGACTGTTTCAGCCTACCCCTTTTCTGAG 1098
Db 1424 TCGACACCCGCTGGGTGGTGTGCGGCCACAGACCTGCCACTGAGAGCTTGGAGAGCAAG 1483
QY 1099 GGTACAGCTTCAAGTGTGCTGTCATCCACCCTGTGTGCTGAACTCCCAAGCTATCAAGGT 1158
Db 1484 GAGGCGCAGCCGACCCACACACCGCTGGATGTGAGTATGCGGCTCCGAGCCGACGTC 1543
QY 1159 CCACACCGCTTATCCAAACATCTGTACACCAAAACCATCTCTGCTCCTCCTATATGAG 1218
Db 1544 CCACCGCTTCTGACCAAGCACTTGTGCTGCCAGTAAGTCTGAGCAGTACAGCGTG 1603
QY 1219 TCAGTAAGTAAGCCCTGCAAGATGAACAGCCCTGAGTACCCAAAGGAAGATGATAGTAG 1278
Db 1604 GACCTGGACAGCTGTGACACCTGTGCGAGCAGCTGCGAGGGGCGCTGAGCAAGAC 1663
QY 1279 AAGAGTCCGACAGT-----GGTACTCTGGGATAAAGAAAAGAACAG 1323
Db 1664 AGCAGTGCACAAAGCCACAAAGGTGCGAAGGAGACAAAGCTGGCAGAGACAAAAGACAG 1723
QY 1324 ATTACAACTTCACTATCACTGTTCGAAAAACAAGAAAGATGAAGAACGACGAGAGTCT 1383
Db 1724 ATTACAGCTTCCCGGTGCTGCGCAAGANTAGCAGAGATGAAGAAAGAGAGAGTCT 1783
QY 1384 CGGATTACAGAGTTACTCCCCACAGGCTTTAAGTCTTTCATGAGCAGCAGTACGAGAAC 1443
Db 1784 CGCATCAAGAGTTACTCCCTTATGCTTCAATTTCTCATGAGCAGCAGCTGGAGAT 1843
QY 1444 GTCTGAAAGTCTCATCAGCAGCTGTGCATCGGAAGAACAGCTAGAAAATGAATGATG 1503
Db 1844 GTCATCAAAACCTACCAAGAGAGGTGAGCGGAGGTGAGCGAGGAGGAGTGGCC 1903
QY 1504 CGGTTGGATTATCTAAGATGCCAGGATCAATGAGAAGATGCTTTCCAGAAAGAG 1563
Db 1904 AAGCTGGGCTCTGTGAGCGGAGCAGGAGGATGAGGAGATGCTTCTACCAGAGGAG 1963
QY 1564 TCTAATCTATATCTCTTAAAGGGCTTAAATGGACAAGTCTATGTTTGAAGATTAAG 1623
Db 1964 TCTAATCTAAGCGCTGAAGAGGGCCAGATGGACAGTCCATGTTTGTGAATCAAG 2023
QY 1624 ACATTAGGAATAGAGCGTTTGTGAAGTCTGTCTAGCAAGAAAAGTCTGATCTAAAGCT 1683
Db 2024 ACTCTAGGCATCGTCTTGGGGAAGTGTGCTGCTTGTAAAGCTGGACACTCAGCT 2083
```

QY	1684	TTGTATGCAACAAAGACTCTTCGAAGAAGAACGCTTCCTCCGAAATCAGGTGGCTCAT	174
Db	2084	CTGTACGCCATGAAGACTCTCAGGAAGAAGGATGCTCTGAACCGGAATCAAGTGGCCCAT	2143
QY	1744	GTCAAAAGCCGAGAGGATATCCTACAGAAAGCCGACAAATGAGTGGTGGTCCGCCCTGTAC	1803
Db	2144	GTCAAGGCTTGAGAGGACATCTCGCTCAAGCAGACAAATGAGTGGTGGTCAAACTCTAC	2203
QY	1804	TACTCTTTCCAGGACAAGGACAACCTTGTTGTGTGAGTGAATACATTCCTCTGGGGGGAT	1863
Db	2204	TACTCTTTCCAGGACAAGGACAAGCCTGTACTTTGTGTGAGTGAATACATACAGGCGGGAT	2263
QY	1864	ATGATGAGCCTATTAAATTAGAAATGGGCATCTTCCTGAAATCTGGCAGCAATTTACATA	1923
Db	2264	ATGATGAGCCTCTGTATCAGGATGCGAGTCTTCCTCTGAGCACCTGGCCCGCTTCTACATT	2323
QY	1924	GCAGAACTTACCTGTGTCAGTTCAAAGTGTTCATAAAATGGTGTTTTATTCATAGAGATTT	1983
Db	2324	GCAGAGTTTGACCCCTGGCCATTGAAAGTGTCCACAAGATGGGCTTTATTCACCGGGACATC	2383
QY	1984	AAACCTTGATTAACATTTTGATTCAACCGTGAATGGCCATATTAATTTGACTGACTTTGGCTTG	2043
Db	2384	AAACCTTGACACACATCTATCGACCTGGATGTCATTAAGCTGACAGATTTTGGGCTC	2443
QY	2044	TGCATCTGGCTTCAGATGGACATGACTCCAAGTACTACCAAGTGGGGATCACCCACGG	2103
Db	2444	TGCACCTGGATTCAGGTGGACTCACAATTTCCAAGTACTACCAAGAAGGGAACCATGAGA	2503
QY	2104	CAAGATAGCATGGATTTCACTAACCAATTTGGGAGATCTCTCCAAATTTGCTGGTGTGGGGAC	2163
Db	2504	CAGGACAGCATGGAGCCGGTGACCTCTGGACGATGTTTCCAACTCTCGCTGTGGAGAC	2563
QY	2164	AGACTCAAGCCACTGGAGCGGAGAGCTGCTCGCCAGCACACAGCGATGTCTAGGCCATTTCT	2223
Db	2564	AGCTTAAAGACCTTGAGCAGAGGCGCGAAGACACACAGAGGTGCTTGGGCACATTTCT	2623
QY	2224	CTGGTTGGGACTCCCAATTTATATTGCACCTGAAAGTCTACTGCGAACAGGATATACACAG	2283
Db	2624	CTTGTGGGCACACCAAAATTTACATGCTCCGGAGGTGCTTCTCCGCCAAAGGTACACGACG	2683
QY	2284	CTGTGTGACTGTGGAGTCTTGGTGTATTCTTTGTGAAATGTTTGGTGGGACAAACCTCCT	2343
Db	2684	CTCTGTGACTGTGGAGCGTGGTGTGATTCCTTTGAGATGCTGGTGGGACGCGCCT	2743
QY	2344	TTCTTTGGCACAAACCCCATTTAGAAACACAAATGAAGGTATATCTCTCGCAAACTCTCTTA	2403
Db	2744	TTCTTTGGCCCCACCCCCACAGACGACGAGTGAAGGTGATCAACTGGGAGACGACGCTG	2803
QY	2404	CACATCCCTCTCAAGCTAAGCTGAGTCTCTGAAGCCTCTGACCTCATTTATCAAACTGTGT	2463
Db	2804	CATATCCCTTACGAGTGAAGCTCAGCCTGAGGCCGAGACCTCATCACGAAGCTGTGC	2863
QY	2464	CGAGGACACAGACGCGCTCGGCAAGAACGGTCTGTATGAGATAAAGGCTCATCTCAATTT	2523
Db	2864	TGCGCGGCTTGACTTGC CGCTTGGGCAGGATGGGCAGATGACCTCAAGGACACACCGTTC	2923
QY	2524	TTTAAACACCATCGATTTCTCTAGTGCATCTGAGACAGACGCTGCTGCTTCATACATCCCTAAA	2583
Db	2924	TTCAACACCATCGACTTTTCCGCTGACATCTCGAAGACGCTGACCCCTAGTCCCCACC	2983
QY	2584	ATCACGATCTCAACAGATACATCCAAATTTCCAGCCTCTGTGATCTGTATAAATTTGAGAC	2643
Db	2984	ATCACGACCCCATGGACACCTTCCAAATTTTGACCCGGTGGATGAAGAAAGCCCTTGAC	3043
QY	2644	GATGGCAGCGAGGAGAAATATCAGTGCACACTCTGAGCGGATGGTATAAAATGGGAAG	2703
Db	3044	GAGGCCAGCGGAGAG- ---AGCGCCAAAGCCCTGGACACGCTTGGCCTCCCCACGACGCAAG	3100
QY	2704	CACCCCGACGACGCTTTCTATGAGTTACCTTTCTCGGAGGTTTTTTGATGACAAATGGCTAC	2763
Db	3101	CATCCAGACGACGCTTCTATGAGTTTACCTTTCCGAGGTTCTTCCGATGACAAAGGCTAT	3160
QY	2764	CCATATAATTATCCAAAGCCTATTGAGTATGAATACATT	2802

Db 3161 CCCTTCGGTGCCCGAAGACCCCTCAGAGCCCGCAGAGAGT 3199

RESULT 13

E38226

LOCUS

DEFINITION Human tumor regulatory gene.

ACCESSION E38226

VERSION E38226.1 GI:18626935

KEYWORDS JP 2000210086-A/1.

SOURCE Homo sapiens.

ORGANISM Homo sapiens

REFERENCE Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

AUTHORS Koga,J., Kono,K. and N.Z.F.

TITLE Human tumor regulatory gene

JOURNAL Patent JP 2000210086-A 1 02-AUG-2000;

COMMENT NIHON CHEMICAL RESEARCH K K

OS Homo sapiens (human)

PN JP 2000210086-A/1

PD 02-AUG-2000

PF 25-JAN-1999 JP 1999016223

PR

PI JUNICHI KOGA,KEIKO KONO,ZOROTARYOFU FIODORU N PC

C12N15/09,A61K31/27,C07K14/47,C07K14/525,C1201/69, PC

G01N33/50,

PC G01N33/50,G01N33/574,G01N33/577//C12P21/08,(C12N15/09 PC

C12R1:91),C12N15/00,

PC A61K37/36,(C12N15/00,C12R1:91)

CC

Key Location/Qualifiers

PH 1.5486

FT source /organism='Homo sapiens (human)'

FT Location/Qualifiers

1.5486

/organism="Homo sapiens"

/db_xref="taxon:9606"

BASE COUNT 1378 a 1452 c 1390 g 1266 t

ORIGIN

Query Match 24.5%; Score 787.2; DB 6; Length 5486;

Best Local Similarity 68.7%; Pred. No. 3.5e-179;

Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps 1;

QY 1220 CAGTAAGTACGCCGTGCAAGATGACACAGCCTAGCTTACCCACGAAGAGATGATAGTGAGA 1279

Db 1999 CAGGCATGTGAGCAGAGAGCCTCCGTGGCGGCCCAACGAGCCGCGGCGGACAAAGAGCC 2058

QY 1280 AGAGTGGCGGACAGTGGTGACTCTGGGGATAAAGAAAGAAACAGATACAACTTCACCTA 1339

Db 2059 GCAAGCGCCACAGGGGGACAAAGCGCGGAAGGATAAAAAGCAGATTACAGACCTCTCCCG 2118

QY 1340 TCACGTGTCGGAAACCAAGAAAGATGAAGAACGAAGAGTCTCGGATTTCAGAGTTACT 1399

Db 2119 TTCGGTCCGCAAAACAGCAGAGACGCAAGAGAGAGAGTACGCGATCAAGAGGTACT 2178

QY 1400 CCCCACAGGCGCTTAAAGTCTTCATGGAGCAGCAGCTAGAGACGCTCCTGAAGTCTCATC 1459

Db 2179 CGCATACGCGCTTAAAGTCTTCATGGAGCAGCAGCTGGAGAATGTCATCAAAACCTACC 2238

QY 1460 AGCAGCGCTCTGCATCGGAACGACGCTAGAAAATGAATGAATGATCGGGTTGGATTATCTC 1519

Db 2239 AGCAGAAGGTTAACCGGAGGCTCGAGCTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTG 2298

QY 1520 AAGATGCCCGCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAAGAGTCTTAATATTTCGTC 1579

Db 2299 AAGCTGAGCAGGACGAGATCCGGAGATCTCTACCAGAAAGAGTCTAATTTACACAGGT 2358

QY 1580 TTAAGAGGGCTAAATGGGCAAGTCTATGTTTGTAAAGATAAAGACATTAGGAATAGGAG 1639

Db 2359 TAAAGAGGGCCAGAGTGAGCAAGTCTATGTTGTCGAAGTCAAAACCCCTGGGATCGGTG 2418

QY 1640 CGTTTGTGTAAGTCTGTCTAGCAAGAAAGTGCATATAAGCTTTGTATGCAACAAGA 1699
 Db 2419 CCTTTGGAGAAGTGTGCTTGTGAAGTGGACACTACGCCCTGTACGCCATGAAGA 2478
 QY 1700 CTCTTCCAAAGAAAGAGTGTCTGCTCCGAATCAGGTGGCTCATGTAAGCGGAGAGGG 1759
 Db 2479 CCCTTAGGAAAGAGTGTCTGCTGAACCGGAATCAGGTGGCCACGTCGAAGCGCGAGAGGG 2538
 QY 1760 ATATCTCTAGCAGAACGCGCAATGAGTGGTGTCCGCTGTACTCTTTCCAGGACA 1819
 Db 2539 ACATCTGCGCGAGGACCAATGAGTGGTGTCAAACTCTACTACTCTTCCAGGACA 2598
 QY 1820 AGGACAACTTGTACTTGTGTGATGAGTACATCTCTGGGGGGATATGATGAGCTTATTA 1879
 Db 2599 AAGACAGCCTGTACTTGTGTGATGAGTACATCTCTGGTGGGACATGATCAGCCTGCTGA 2658
 QY 1880 TTAGATGGGCATCTTCTTCCGAAATCTGGCAGATCTACATAGCAGAACTTACCTGTG 1939
 Db 2659 TCCGGATGGAGGTCTTCCCTGAGCACTGCGCCGGTCTACATCGCAGAGCTGACTTTGG 2718
 QY 1940 CAGTTGAACTGTTCATAAAATGGTTTTATTATCATAGAGATATTAAACCTGATAACATTT 1999
 Db 2719 CCATTGAGAGTGTCCACAGATGGCTTCATCCACGACATCAACGCTGATACATTT 2778
 QY 2000 TGATTGACCTGATGGCCATTTAAATTTGACTGACTTTGGCTTGTGCACCTGGCTTCAGAT 2059
 Db 2779 TGATAGATCTGGATGTGTCATTAATCAACATCAAGATTTGGGCTCTGCACCTGGCTCAGGT 2838
 QY 2060 GGACACATGACTCCAGTACTACAGAGTGGGATCACCACCGCAAGATAGCATGGATT 2119
 Db 2839 GGACTCACAATTTCCAAATATTACCAAGAGGAGGACCTGTGCAGACGAGCAGACGAGC 2898
 QY 2120 TCAGTAACGAATGGGAGATCTTCCAAATTTGCTGGTGGGACAGCTCAAGCCACTGG 2179
 Db 2899 CAGGACCTCTGGATGATGTCTTAAGTGTGCTGGTGGGACAGCTGGAAGACCTAG 2958
 QY 2180 AGCGGAGAGTCTGCCAGCACCAGCGATGTCTAGCCCATCTCTGGTTGGGACTCCCA 2239
 Db 2959 AGCAGAGGCGCGGAAGCAGCAGCAGAGGTGCTGCGACATTCACCTGGTGGGACTCCAA 3018
 QY 2240 ATTATATTGCACCTGAGTGTCTACTCGGACAGGATATACACAGCTGTGTGACTGTGA 2299
 Db 3019 ACTACATGCGACCCGAGGTGCTCTCCGAAAGGGTACACTCAACTCTGTGACTGTGA 3078
 QY 2300 GTGTTGGTGTATTCTTTGTGAATGTGGTGGCAACTCTCTTTCTTGGCAACAACC 2359
 Db 3079 GTGTTGGAGTGTCTTCTCGAGATGCTGTGGGCGAGCGCCCTTTTGGCACCTACTC 3138
 QY 2360 CATTAGAAACAAATGAAGGTATCATCTGGCAAACTTCTCTACACATCCCTTCCTCAAG 2419
 Db 3139 CCACAGAAACCCAGCTGAAGGTATCAACTGGGAGAACACGCTCCACATTCAGACCCAGG 3198
 QY 2420 CTAAGCTGAGTCTGAAGCTCTGACCTCATATCAAACTGTGTGAGGACCAAGAGACC 2479
 Db 3199 TGAAGCTGAGCCTGTAGGCGCAGGACCTCATCAACCAAGTGTGCTCTCCGACAGACCC 3258
 QY 2480 GCCTCGGCAAGACGCTGTGATGAGATAAGGCTCATCCATTTTAAAGACCATCGATT 2539
 Db 3259 GCCTGGGCGGAATGGGCGGATGACCTGAAGGCCACCCCTTCTTCAGGCGCATTTGACT 3318
 QY 2540 TCTCTAGTATCTGACAGCAGCTGTCTTCTATACATCCCTAAATCAGGCACTCCAAACAG 2599
 Db 3319 TCTCAGTGACATCCGGAAGCAGCCAGCCCTTACGTTCCCAACCATCAGCCACCCCATGG 3378
 QY 2600 ATACATCCAAATTCGACCTGTGTGATCTCTGATTAATTTGGAGCGATGCGAGGAGGAG 2659
 Db 3379 ACACCTCGAATTCGACCCCGTAGATGAAGAAGCCCTTGGAAACGATGCCAGCAGCAGC 2719
 QY 2660 AAAATATCAGTGACACTCTGAGCGGATGTATAAATGGGAAGCAGCCCGGAGCAGCTT 2719
 Db 3436 GTAGCACCAGGCGCTGGGACACACTCACCTCGCCCAATTAACAGCATCTTGAGCAGCAT 3495

QY 2720 TCTATGAGTTCACCTTCGGAGGTTTTTTTGTATGACAATGCTACCCATATAATATCCAA 2779
 Db 3496 TTTACGAATTCACCTTCGGAAGGTTCTTTGATGACAATGCTACCCCTTCGATGCCAA 3555
 QY 2780 AGCCTATTGATGATGATACATTCATTCATTCACAGGCTCAGA 2819
 Db 3556 AGCCTTCAGGAGCAGAAGTTCACAGGCTGAGAGCTCAGA 3595
 RESULT 14
 E38227
 LOCUS Human tumor regulatory gene.
 DEFINITION E38227
 ACCESSION E38227
 VERSION E38227.1 GI:18626936
 KEYWORDS JP 2000210086-A/2.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 Koga, J., Kono, K. and N, Z. F.
 Human tumor regulatory gene
 Patent: JP 2000210086-A 2 02-AUG-2000;
 NIHON CHEMICAL RESEARCH K K
 OS Homo sapiens (human)
 PN JP 2000210086-A/2
 PD 02-AUG-2000
 PF 25-JAN-1999 JP 1999016223
 PR
 PI JUNICHI KOGA, KEIKO KONO, ZOROTARYOFU FIODORU N PC
 C12N15/09, A61K31/00, A61K38/27, C07K14/47, C07K14/525, C12Q1/68, PC
 G01N33/50,
 PC G01N33/50, G01N33/57, G01N33/577, C12P21/08, (C12N15/09, PC
 C12R1:91), C12N15/00,
 PC A61K37/36, (C12N15/00, C12R1:91)
 CC
 FT source
 Location/Qualifiers
 1. 5486
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 BASE COUNT 1380 a 1451 c 1388 g 1267 t
 ORIGIN
 Query Match 24.5%; Score 787.2; DB 6; Length 5486;
 Best Local Similarity 68.7%; Pred. No. 3.5e-179;
 Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps 1;
 QY 1220 CAGTAAGTAAGCCTCGAAAGATGAACAGCTAGCTTACCCAGGAGGATGATAGTGAGA 1279
 Db 1999 CAGGCATGAGCAGAGACCTCCGTCGCGGCCCAACGAGGCCGCGGCGGCAAGAGCC 2058
 QY 1280 AGAGTCGGGACAGTGTGACTCTCTGGGGATAAAGAAAGAAACAGATTACAACTTCACTA 1339
 Db 2059 GCAGAGCCCAAGGGGACAAAGCGGAAGGATAAAGCAGATTCAGACCTCTCCCG 2118
 QY 1340 TCAGTGTTCGGAACAAAGAAAGATGAAGAACGAGAGTCTCGGATTCAGAGTTACT 1399
 Db 2119 TTTCCGCTCGCAAAACACAGCAGAGCAAGAGAGAGAGTCCACATCAAGAGCTACT 2178
 QY 1400 CCCCACAGCCCTTAAAGTTCTTCATGGACGACGCTAGAGACGCTCTTCAAGTCTCATC 1459
 Db 2179 CGCCATACCCCTTTAAGTTCTTCATGGACGACGACCGTGGAGATGTCTCAAAACCTACC 2238
 QY 1460 AGCAGCGCTCTGCATCGAAAGAGCAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTC 1519
 Db 2239 AGCAGAGGTTAAACCGGAGGCTGCAGCTGGAGCAGAAATGGCCAAAGCTGGACTCTGTG 2298
 QY 1520 AAGATGCCCGAGTCAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCCTGC 1579

Db 2299 AAGCTGACGAGGAGGATCGGGAAGATCCTCTACCAGAAAGAGTCTAATTACAACAGGT 2358
Qy 1580 TTAAGAGGCTAAAATGACAAAGTCTATGTTGTAAGATAAAGACATTAGGAATAGGAG 1639
Db 2359 TAAGAGGCGCAAGTGAAGATCTATGTTGTAAGATCAAAACCTCGGGATCGGTG 2418
Qy 1640 CGTTTGGTGAAGTCGTGTAGCAAGAAAGTGTGATCTAAAGCTTTGTATCAACAAGAA 1699
Db 2419 CTTTGGAGAGAGTGTCCTTGTGTAAGGTGGACACTCACGCCCTGTACGCCATGAAGA 2478
Qy 1700 CTCCTCGAAGAAAGACGCTCTGCTCCGAATCAAGTGGCTCATGTGAAGCGGAGAGGG 1759
Db 2479 CCCTAAGGAAAGAGATGCTCTGAACCGGAATCAGTGGCCACAGCTCAAGGCCGAGAGG 2538
Qy 1760 ATATCTAGCAGAGCCGACAAATGAGTGGTGGTCCGCTGTACTACTCTTTCCAGGACA 1819
Db 2539 ACATCTGCGCGAGGACAGCAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGACA 2598
Qy 1820 AGGACAACCTTTGATCTGATGAGTACATCTCTGGGGGATATGATGAGCCCTATTAA 1879
Db 2599 AAGACAGCCTGTACTTTGTGATGAGTACATCTCTGGTGGGACATGATGAGCCTGTGA 2658
Qy 1880 TTAGATGGGATCTTTCTGAAATCTGGCAGATTTCTATAGAGATTAATAACTGATACATTT 1939
Db 2659 TCCGATGGAGGCTCTCTCCCTGAGCAGCTGGCCGCTTCTATCGCAGAGTGAATTTGG 2718
Qy 1940 CAGTTGAAAGTCTTCAAAAATGGTTTTATTATAGAGATTAATAACTGATACATTT 1999
Db 2719 CCATGAGAGTGTCCCAAGATGGCTTCATCCAGGAGACATCAAGCCTGATACATTT 2778
Qy 2000 TGATTGACCGTGTGGCCATTAATAATGACTGATTTGGTGTGTCAGCTGGCTTCAGAT 2059
Db 2779 TGATGATCTGGTGTGATCAATTAATCAAGATTTTCGGCTCTGCACTGGTTTCAGT 2838
Qy 2060 GGACATGACTCCAGTACTACAGAGTGGGATCAGCCAGCGAATGATGATGATGATTT 2119
Db 2839 GGACTCAATATCCCAATATTACCAAGAGGAGCCATGTGAGCAGGACGATGAGG 2898
Qy 2120 TCAGTAAGCAATGGGAGATCCTTCAATGTCGGTGTGGGACAGACTGAAGCCACTGG 2179
Db 2899 CCAGGACCTCTGGATGATGTCTACTGTGCTGTGGGACAGGCTGAAGACCTTAG 2958
Qy 2180 AGCGGAGAGTGTGCGGACGACGATGCTAGCCCATCTCTGTTGGGACTCCCA 2239
Db 2959 AGCAGAGGCGGGAAGCAGCAGCAGAGGTGCTGGCAGATTCACCTGTGGGACTCCA 3018
Qy 2240 ATTATATGCACTGAAGTGTCTGCAAGAGGATATACAGAGTGTGCTGACTGGTGA 2299
Db 3019 ACTACATGCAAGGAGTGTCTCTCGCAAGAGGTGCTGGCAGATTCACCTGTGACTGGTGA 3078
Qy 2300 GTGTTGGTGTATTCTTTGTAATGTTGGTGGGACACCTCTCTTTGGCAGCAAGCC 2359
Db 3079 GTGTTGGATGATCTCTTCAGATGCTGTGGGCGAGCCGCTTTTGGCAGCTACTC 3138
Qy 2360 CATTAGAACCAATGAAGGTTATCATCTGGCAAACTCTCTACACATCCCTCTCTCAAG 2419
Db 3139 CCACAGAAACCCAGCTGAAGTGTCACTGGGAGAACACCTCCACATTCACGCCAGG 3198
Qy 2420 CTAGCTGAGTCTCTGAGCCCTGACCTCATATCAACTGTGTGCGAGGACAGAGACC 2479
Db 3199 TGAAGCTGAGCCCTGAGCCAGGAGCCTCATCAACCACTGTGCTGCTCCGAGCACC 3258
Qy 2480 GCCTCGGCAAGAGGCTCTGATGAGATGAAGGCTCATCCATTTTTTAAGACCATCGATT 2539
Db 3259 GCCTGGGCGAATGGGCGCATGACCTGAAGGCCACCCCTCTTCTAGCGCCATTGACT 3318
Qy 2540 TCTCTAGTGTCTGAGACAGAGTGTCTTCTATCATATCCCTTAAATACCGATCCCAAG 2599
Db 3319 TCTCAGTGATCATCCGGAAGCAGCCAGCCCTAGTCTTCCACCATCAGCCACCCCATG 3378
Qy 2600 ATACATCAATTTGACACCTGTGATCTGTATTAATTTGAGGAGGATGGAGCGAGGAGG 2659
Db 3379 ACACCTCAATTTGACACCCCGTAGATGAAGAGCCCTTGAAGAGTCCAGC---GAAG 3435

Qy 2650 AAAATATCAGTGACACTCTGAGCGGATGGTATAAAATGGAAGCACCCGAGCAGCTT 2719
Db 3436 GTAGCACCAAGCCCTGGGACACACTCACCTGCCCAATAACAAGCATCTGAGCAGCAT 3495
Qy 2720 TCTATGAGTTCACCTTCGGAGGTTTTTTTGTATGACAAATGGCTTACCATATATAATTATCAA 2779
Db 3496 TTTAGGAATTCACCTTCGGAAGTTCTTTGTATGACAAATGGCTTACCCCTTCGATGCCAA 3555
Qy 2780 AGCCTATTGATATCAATACATTTCATTCACAGGGCTCAGA 2819
Db 3556 AGCCTCAGGACGACGAAGCTTCACAGGCTGAGAGCTCAGA 3595

RESULT 15
LOCUS AB028019 3412 bp mRNA linear PRI 09-MAR-2000
DEFINITION Homo sapiens HSLATS2 mRNA for large tumor suppressor 2, partial cds.
ACCESSION AB028019
VERSION AB028019.1 GI:7212789
KEYWORDS large tumor suppressor 2; HsLats2.
SOURCE Homo sapiens cDNA to mRNA.
ORGANISM Homo sapiens
REFERENCE 1 (sites)
AUTHORS Yabuta,N., Fujii,T., Copeland,N.G., Gilbert,D.J., Jenkins,N.A., Nishiguchi,H., Endo,Y., Toji,S., Tanaka,H., Nishimune,Y. and Nishiguchi,H.
TITLE Structure, expression, and chromosome mapping of LATS2, a mammalian homologue of the Drosophila tumor suppressor gene lats/warts
JOURNAL Genomics 63 (2), 263-270 (2000)
MEDLINE 20139436
REFERENCE 2 (bases 1 to 3412),
AUTHORS Nojima,H. and Nishiguchi,H.
TITLE Direct Submission
JOURNAL Submitted (27-MAY-1999) Hiroshi Nojima, Research Institute for Microbial Diseases, Osaka University, Department of Molecular Genetics, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
(E-mail:inojima@biken.osaka-u.ac.jp, Tel:81-6-6875-3980, Fax:81-6-6875-5192)

FEATURES
source
1..3412
Location/Qualifiers
/organism="Homo sapiens"
/db_xref="taxon:9606"
1..3412
/gene="HsLATS2"
/gene="HsLATS2"
/codon_start=3
/product="large tumor suppressor 2"
/protein_id="BA02381.1"
/db_xref="GI:7212790"
/translation="NSDTSLDKAVLGSKDATSSKQMRATPKFGYOKALREIRYSL
PFANESGTSAAAEVNRQMLQELNAGCQEMAGRALKQTSRSTEALEYISRMGYLD
PRNEQIVRVIKOTSPGKGLMPTVTRRPSFGTGDSPFASVHOLSTGYEPGSPFADGP
TALGEMPRVYDVLFGVGPCHQHPKPGYASVGAAGAFHPLOGAHYRGPHELL
VPEPLGYVQRSPFSFKTPTTGGTASLTPTGQGGPPGAGLAFPPAAGLYVPHEH
HKQAGVAHQHLVILGSRQSFADSPQSLTPTSRNSLNDVLYLSSTSVQOQPAATL
ARRDLSQGLGEPRAHVAFRPDCVPVSRSTNSFNHQPPGPPGKAEPSPAPNTVT
VSAAHLLHPVKSVRVPOTAVGSHPAWVPAAPAPAPAPAPAAEGJDAKEEHA
LALGAGAFPLDVEYGGPDRSARLPTRSTCCCAKSKQYDLDLSLCAGMEOSLRAGNE
PEGDKSRKSAKDGKDKKQIQTSVPVYKNSRDEKRESIKSYSPYAFKPFMEQ
HVENVIKTYQKVNRLQLOEMAKGLCEAQEMKRLILQRESNTRUKRAKMS
MFVKITGLTGAFGEVCLACKDTHALYAMKTLRKKLNRLNNOVAHVKAERDILAEAD
NEWVKLYYFQDKDLSLYFVMDYIPGDDMSLLIRMEVFHEHARFYIAETLAIESV
HKMGFIHRDIKPNILDLGHIKTLDFGLCTGFRWTHNSKYOKGSHVRODSMEPSD
LWDDVSNRCGDRDLKLEQARQKHORSLAHSINGTNPYIAPEVLLRKGTLCCDWS
VGYLLFEMLYGQPPFLAPTETETOLKVINVENTLHIIPAQVKLSPEARDLITKCCSD
HLGRNGADDLKAHPFFSAIDFSSDIRKQPPAPVYPTISHPMDSNFDVDEESPNDA
SEGSKAWDLTISPNKKHFAFYETFRFFDDNGYFPFRCPKPSGAESAQSSDLE
SSDLVDQTEGCPVYV"

polYA_signal 3364 3369
BASE COUNT 801 a 1066 c 975 g 570 t
ORIGIN

Query Match 24.5%; Score 785.6; DB 9; Length 3412;
Best Local Similarity 68.6%; Pred No. 8.4e-179;
Matches 1098; Conservative 0; Mismatches 499; Indels 3; Gaps 1;

QY 1220 CAGTAAGTAAAGCCCTCAAGATGAACAGCTAGCTTTACCCAAAGGAAGATGATAGTGAGA 1279
DB 1489 CAGGCATGGAGCAGAGCCCTCGGTGCGGGCCCAACAGCAGCCGAGCGCGGCAAGAGCC 1548
QY 1280 AGAGTCGGACAGCTGTGACTCTGGGGATTAAGAAAAGAAACAGATTCAACTTCACCTA 1339
DB 1549 GCAAAAGCCCAAGGGGCAAGAGGAAAGGATTAAGAGAGATTCAGACCTCTCCCG 1608
QY 1340 TCACTGTCGGAAGAAACAGAAAGATGAAGAACGAAGAGTCTCGGATTCAGAGTTACT 1399
DB 1609 TTCCTCCGCAAAACAGCAGAGACGAAGAGAGAGAGTCAACATCAAGAGCTACT 1668
QY 1400 CCCACAGCCCTTTAAGTCTTCATGGAGCAGCAGCTAGAGACGCTCCTGGAAGTCTCATC 1459
DB 1669 CGCATACGCCCTTTAAGTCTTCATGGAGCAGCAGCTGGAGATGTCAATAAACCTTACC 1728
QY 1460 AGCAGCGTCTGCATCGGAAGAACAGCAGTGAATAATGAATGATCGGGTTGGATTATCTC 1519
DB 1729 AGCAAGAGGTAAACCGGAGCTCGAGCTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTG 1788
QY 1520 AAGATGCCAGGATCAAAATCAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCGTC 1579
DB 1789 AAGCTGAGCAGGAGCAGATCGGGAAGATCCTCTACCAGAAAGTCTAATTAACACAGGT 1848
QY 1580 TTAAGAGGCTAAATGGCAAGCTATGTTCTTAAGATAAAGACATTAGGAATAGGAG 1639
DB 1849 TAAAGAGGCAAGATGGCAAGCTATGTTCTCAAGATCAAAACCCCTGGGGATCGGTG 1908
QY 1640 CGTTTGGTGAAGTCTGCTAGCAAGAAAGTCTGATATCAAGCTTTGTATGCAACAAAGA 1699
DB 1909 CCTTTGGAGAGTGTGCTTGTGTTGTAAGTGGACACTCAGCCCTGTAGGCCATGAAGA 1968
QY 1700 CTCTTGGAAAGAGAGTGTCTCGGAAATCAGTGGCTCATGTGAAGCGGAGAGGG 1759
DB 1969 CCCTAAGGAAAGAGGTGCTTCTGAACCGGAATCAGTGGGCCACGCTCAAGGCCGAGGG 2028
QY 1760 ATATCTAGCAGAGCGCAAGTGAAGTGGTGGTGGCTGTACTCTTTCCAGGACA 1819
DB 2029 ACATCTGGCCGAGGAGACAAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2088
QY 1820 AGGACAACTTGTACTTTGTGATGGACTACATCTCTGGGGGATATGATGAGCCTATFAA 1879
DB 2089 AAGACACCTGTACTTTGTGATGGACTACATCTCTGGTGGGACATGATGAGCCTGTGTA 2148
QY 1880 TTAGATGGGCATCTTTCTTGAATAATCTGCAGATTTCTACATAGCAGAACTTACCTGTG 1939
DB 2149 TCCGGATGGAGGTCTTCCCTTGAGCACCCTGCGCGGTTCTACATCGCAGAGCTGACTTTG 2208
QY 1940 CAGTTGAAGCTGTCAATAAATGGTTTTATTCATAGATATTAACCTGTATAAATTT 1999
DB 2209 CCATTGAGAGTGTCCCAAGATGGCTTTATCCACCAGACATCAAGCCTGTATAAATTT 2268
QY 2000 TGATTGACCGTGTAGGCCATATTAAATTTGACTGACTTTGGCTTGTGCACCTGGCTTCAGAT 2059
DB 2269 TGATAGATCTGGATGGTGCACATTAACATCAGATTTTCGGCCTCTGCACCTGGGTTCAGGT 2328
QY 2060 GGACACATGACTCCAAGTACTACCAGAGTGGGATCACCACCGCAGATAGCATGATT 2119
DB 2329 GGACTCAAAATTCACAAATATTAACCAAGAGGAGCCATGTACAGAGGAGCATGGAGC 2388
QY 2120 TCAGTAACGAATGGGAGATCCCTTCCAAATGTCCGTTGGGGACAGACTGAAGCCACTGG 2179
DB 2389 CCACGACCTCTGGATGATGTGTCTAACTGTCTGCTGCTGGGACAGGCTGAAGACCCCTAG 2448

QY 2180 AGCGGAGAGCTGCTCGCCAGCACCAGCAGATGCTAGCCCAATTTCTGTGTGGACTCCCA 2239
DB 2449 AGCAGAGGCGCGGAGCAGCAGCAGAGGTGCTGSCACATTTCACTGTGGGACTCCAA 2508
QY 2240 ATTATATTGCACCTGAAGTGTCTACTGCGGAACAGGATATACACAGCTGTGTGCTGGA 2299
DB 2509 ACTACATCGACCCGAGGTGCTCTCCGCAAGGGGTACACTCAACTCTGTGTGCTGGA 2568
QY 2300 GTGTTGGTCTTATCTTCTGTAATCTGCTGGGACAACTCTCTTTTGGGCACAAACC 2359
DB 2569 GTGTTGGATGATCTCTTCGAGATGCTGTGGGAGCGGCCCTTTTGGGACCTACTC 2628
QY 2360 CATTAGAAACACAAATGAAGTTATCATCTGGCAAACTTTCTACACATCTCCCTCAAG 2419
DB 2629 CCACAGAAACCCAGCTGAAGTGTATCACTGGGAGAACAGCTCCACATTTCCAGCCAGG 2688
QY 2420 CTAAGTGTAGTCTGAGGCTCTGACCTCATTTATCAACTGTGTGAGGACCAAGAGACC 2479
DB 2689 TGAAGCTGAGCCTTGAGGCCAGGACCTCATCACCAGCTGTGCTCTCCGAGACCACT 2748
QY 2480 GCCTGGCAAGAACGCTGCTGATGATGAATAAGGCTCATCATTTTAAAGACCATCGATT 2539
DB 2749 GCCTGGGCGGAATGGGCGCGATGACCTGAAGGCCACCCCTTCTTCAGGCCATTTGACT 2808
QY 2540 TCTCTAGTGTATGAGACAGCAGTCTGCTTATACATCCCTAAATTCACGATCCCAACAG 2599
DB 2809 TCTCCAGTGACATCGGGAAGCAGCAGCCAGCCCTTACCTCCACCATCAGCCACCCATGG 2868
QY 2600 ATACATCCAAATTCGACCCCTGTTGATCTCTGATTAATTTGGAGCGATGGCAGGAGGAG 2659
DB 2869 ACACCTCGAATTCGACCCCTGATGAAGAAAGCCCTTGGAAACGATGCCAGC 2925
QY 2660 AAAATATCAGTGACACTCTGAGCGGATGCTATAAAATGGGAAGCACCCCGAGCAGCTT 2719
DB 2926 GTAGCAGCAGGCTCGGACACACTCACCTCGCCCAATTAACAGCATCTGAGCAGCAT 2985
QY 2720 TCTATGAGTTCACCTTTGCGAGGTTTTTGTATGACAAATGGCTACCATATTAATTCGA 2779
DB 2986 TTTACGAATTCACCTTTCCGAAGGTTCTTTGTATGACAATGGCTACCCCTTTTCGATGCCAA 3045
QY 2780 AGCCTATTGAGTATGAATACATTTCACAGGGCTCAGA 2819
DB 3046 AGCCTTCAGGACGAGAGCTTCACAGGCTCAGAGCTCAGA 3085

Search completed: January 16, 2003, 18:29:01
Job time : 6446.88 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:07 ; Search time 409.821 Seconds
(without alignments)
17655.667 Million cell updates/sec

Title: US-09-763-334-3

Perfect score: 3213

Sequence: 1 gtgcaacattcaattacgg.....cccccaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N_Geneseq_101002.*

- 1: /SID22/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
- 2: /SID22/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
- 3: /SID22/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
- 4: /SID22/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
- 5: /SID22/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
- 6: /SID22/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
- 7: /SID22/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
- 8: /SID22/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
- 9: /SID22/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
- 10: /SID22/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
- 11: /SID22/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
- 12: /SID22/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
- 13: /SID22/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
- 14: /SID22/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
- 15: /SID22/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
- 16: /SID22/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
- 17: /SID22/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
- 18: /SID22/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
- 19: /SID22/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
- 20: /SID22/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
- 21: /SID22/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
- 22: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
- 23: /SID22/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
- 24: /SID22/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3213	100.0	3213	21	AAZ51506
2	3209.8	99.9	3213	17	AAZ51506
3	2483	77.3	3984	17	AAZ51506
4	2483	77.3	3984	21	AAZ51506
5	2481.4	77.2	7382	20	AAZ51506
6	1946.8	60.6	2442	20	AAZ51506
7	1126.8	35.1	1374	20	AAZ51506
8	1012	31.5	1357	22	ABA08740
9	812.6	25.3	3155	17	AAZ51506

10	812.6	25.3	3155	21	AAZ51507
11	787.2	24.5	5276	20	AAZ51507
12	787.2	24.5	5486	21	AAZ51507
13	787.2	24.5	5486	21	AAZ51507
14	785.6	24.5	1961	21	AAZ51507
15	785.6	24.5	3533	22	AAZ51507
16	730.6	22.7	1912	23	AAZ51507
17	713.8	22.2	1498	21	AAZ51507
18	681	21.2	2043	22	AAZ51507
19	598.2	18.6	795	22	AAZ51507
20	521.2	16.2	5720	17	AAZ51507
21	521.2	16.2	5720	21	AAZ51507
22	519.6	16.2	3319	23	AAZ51507
23	518.6	16.1	638	21	AAZ51507
24	492.4	15.3	1501	22	AAZ51507
25	468.4	14.6	582	20	AAZ51507
26	376.6	11.7	678	21	AAZ51507
27	371.8	11.6	676	22	AAZ51507
28	371.8	11.6	676	23	AAZ51507
29	252.4	7.9	11187	23	AAZ51507
30	248.4	7.7	1935	20	AAZ51507
31	248.4	7.7	4983	22	AAZ51507
32	244.4	7.6	3018	17	AAZ51507
33	244.4	7.6	3018	24	AAZ51507
34	244.4	7.6	3800	22	AAZ51507
35	226.4	7.0	1710	21	AAZ51507
36	226.4	7.0	2003	21	AAZ51507
37	221.6	6.9	1894	21	AAZ51507
38	220	6.8	1922	21	AAZ51507
39	211.6	6.6	2001	21	AAZ51507
40	210.2	6.5	2255	23	AAZ51507
41	198.6	6.2	2101	17	AAZ51507
42	179.8	5.6	2160	22	AAZ51507
43	175.6	5.5	734	23	AAZ51507
44	175.2	5.5	734	22	AAZ51507
45	175.2	5.5	734	22	AAZ51507

ALIGNMENTS

RESULT 1
AAZ51506
ID AAZ51506 standard; DNA; 3213 BP.

XX AAZ51506;

XX 21-JUN-2000 (first entry)

XX Mouse Lats (large tumour suppressor) DNA.

XX Mouse: Lats; large tumour suppressor; cytostatic; vulnary;
XX cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
XX treatment; prevention; screening; cancer; skin; ovarian tumour;
XX soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
XX LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
XX dysplasia; degenerative disorder; growth deficiency; physical trauma;
XX hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX Mus musculus.

XX Key Location/Qualifiers
XX CDS 1..2889
XX /*tag= a
XX /product= "Lats protein"
XX /partial

XX WO200010602-A1.

XX 02-MAR-2000.

XX 18-AUG-1999; 99WO-US19068.

XX

Db 1561 GAGTCTAATATATCGCTTTAAAGGGCTAAATGGACAAAGTCTATGTTTGAAGATA 1620
Qy 1621 AAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTGCATACATAA 1680
Db 1621 AAGACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTGCATACATAA 1680
Qy 1681 GCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGTGCGCT 1740
Db 1681 GCTTTGTATGCAACAAAGACTCTTCGAAAGAAAGACGTTCTGCTCCGAAATCAGTGCGCT 1740
Qy 1741 CATGTGAAGCGGAGGGATATCCTAGCAGAACCGACATGAGTGGTGGTCCGCGTG 1800
Db 1741 CATGTGAAGCGGAGGGATATCCTAGCAGAACCGACATGAGTGGTGGTCCGCGTG 1800
Qy 1801 TACTACTCTTTCCAGGACAAAGCAACTTGTACTTTGTGATGGACTACATTCCTGGGGGG 1860
Db 1801 TACTACTCTTTCCAGGACAAAGCAACTTGTACTTTGTGATGGACTACATTCCTGGGGGG 1860
Qy 1861 GATATGATGAGCCATTAAATAGAAATGGGCATCTTCTGAAATCTGGCAGCATTTCTAC 1920
Db 1861 GATATGATGAGCCATTAAATAGAAATGGGCATCTTCTGAAATCTGGCAGCATTTCTAC 1920
Qy 1921 ATACGAGAACTTACTGTGACGTTGAAGTGTTCATAAATGGGTTTATTCATAGAGAT 1980
Db 1921 ATACGAGAACTTACTGTGACGTTGAAGTGTTCATAAATGGGTTTATTCATAGAGAT 1980
Qy 1981 ATAAACCTGTAAATCTTTGATGACCGTATGCCATATTAATTTGACTGACTTTGGC 2040
Db 1981 ATAAACCTGTAAATCTTTGATGACCGTATGCCATATTAATTTGACTGACTTTGGC 2040
Qy 2041 TTGTGCACTGGCTTCAGATGACACATGACTCCAAAGTACTACCAGAGTGGGATCACCCA 2100
Db 2041 TTGTGCACTGGCTTCAGATGACACATGACTCCAAAGTACTACCAGAGTGGGATCACCCA 2100
Qy 2101 CGGCAAGATAGCATGGATTTTCAGTAAGCATGGGAGATCTTCCATTTGTCGGTGGG 2160
Db 2101 CGGCAAGATAGCATGGATTTTCAGTAAGCATGGGAGATCTTCCATTTGTCGGTGGG 2160
Qy 2161 GACAGACTGAAGCCACTGGAGCGAGAGTCTGCTCCAGCAGCAGCGATGCTAGCCCAT 2220
Db 2161 GACAGACTGAAGCCACTGGAGCGAGAGTCTGCTCCAGCAGCAGCGATGCTAGCCCAT 2220
Qy 2221 TCTCTGTTGGAGTCCCAATTTATATTGCACCTGAAGTGTCTGCGAACAGGATATACA 2280
Db 2221 TCTCTGTTGGAGTCCCAATTTATATTGCACCTGAAGTGTCTGCGAACAGGATATACA 2280
Qy 2281 CAGCTGTGTACTGGTGGAGTGTGGTGTATTCTTTGTGAAATGTTGGTGGGACACCT 2340
Db 2281 CAGCTGTGTACTGGTGGAGTGTGGTGTATTCTTTGTGAAATGTTGGTGGGACACCT 2340
Qy 2341 CTTTCTTGGCACAACCCCAATTAAGAACACAAATGAAGTGTATCATCTGGCAAACTTCT 2400
Db 2341 CTTTCTTGGCACAACCCCAATTAAGAACACAAATGAAGTGTATCATCTGGCAAACTTCT 2400
Qy 2401 CTACATATCCCTCTCAAGCTTAAGCTGAGTCTGAAGCCCTGACCTCATTTATCAAACTG 2460
Db 2401 CTACATATCCCTCTCAAGCTTAAGCTGAGTCTGAAGCCCTGACCTCATTTATCAAACTG 2460
Qy 2461 TGTCGAGGACCAAGAGCCGCTCGGCAAGAACGGTGTGATGAGATAAAGGCTCATCCA 2520
Db 2461 TGTCGAGGACCAAGAGCCGCTCGGCAAGAACGGTGTGATGAGATAAAGGCTCATCCA 2520
Qy 2521 TTTTAAAGACCATCGATTTCTCTAGTGTCTGAGACGAGTGTGCTTTCATACATCCCT 2580
Db 2521 TTTTAAAGACCATCGATTTCTCTAGTGTCTGAGACGAGTGTGCTTTCATACATCCCT 2580
Qy 2581 AAAATCAGCATCCAAAGATACATCCAAATTTTCAGCCCTGTGTATCTGATFAAATTTGG 2640
Db 2581 AAAATCAGCATCCAAAGATACATCCAAATTTTCAGCCCTGTGTATCTGATFAAATTTGG 2640
Qy 2641 AGCGATGCGAGGAGGAAATATCAGTGACACTCTGACGCGATGATATAAATGGG 2700
Db 2641 AGCGATGCGAGGAGGAAATATCAGTGACACTCTGACGCGATGATATAAATGGG 2700

Qy 2701 AAGCACCCCGAGCAGCGTTTCTATGAGTTCACTTTTCGAGGTTTTTTTGATGACAATGGC 2760
Db 2701 AAGCACCCCGAGCAGCGTTTCTATGAGTTCACTTTTCGAGGTTTTTTTGATGACAATGGC 2760
Qy 2761 TACCCATATATATATCCAAAGCCCTATTGAGTATGAATACATTCATTCACAGGGCTCAGAA 2820
Db 2761 TACCCATATATATATCCAAAGCCCTATTGAGTATGAATACATTCATTCACAGGGCTCAGAA 2820
Qy 2821 CAACAGTCTGATGAGATGATCAACACACAGCTCCGATGGAACAAACCGAGATCTAGTG 2880
Db 2821 CAACAGTCTGATGAGATGATCAACACACAGCTCCGATGGAACAAACCGAGATCTAGTG 2880
Qy 2881 TATGTTTAAATAAAGTAGGAGATCATTTGTAAGAAATTTGCAAGAGCCCTGAAGTGCAGGGT 2940
Db 2881 TATGTTTAAATAAAGTAGGAGATCATTTGTAAGAAATTTGCAAGAGCCCTGAAGTGCAGGGT 2940
Qy 2941 TTTTGAAGTTTTTGAGAAAATTTATGCAAAATGTGACAGAGTTTGTGCTCTGTGACAATA 3000
Db 2941 TTTTGAAGTTTTTGAGAAAATTTATGCAAAATGTGACAGAGTTTGTGCTCTGTGACAATA 3000
Qy 3001 TTTTATTTTCTTAAGTTATGGAAATTTGTTTAAATGTTTAAATTTATTTCCACCCTTTAA 3060
Db 3001 TTTTATTTTCTTAAGTTATGGAAATTTGTTTAAATGTTTAAATTTATTTCCACCCTTTAA 3060
Qy 3061 TTCAGTAAATTTAGAAAATTTGTTTAAAGAAAGTAAATTTAACTGAGTATTATAGTC 3120
Db 3061 TTCAGTAAATTTAGAAAATTTGTTTAAAGAAAGTAAATTTAACTGAGTATTATAGTC 3120
Qy 3121 AATTTCTTGGTAAAGTAAAGTAAAGAGAGAGCCCTGGTATCTTTTGTATATATAATAA 3180
Db 3121 AATTTCTTGGTAAAGTAAAGTAAAGAGAGAGCCCTGGTATCTTTTGTATATATAATAA 3180
Qy 3181 ATAATTTTAAATCCCAAAAAAATAAAAAA 3213
Db 3181 ATAATTTTAAATCCCAAAAAAATAAAAAA 3213

RESULT 2

AA42119
ID AA42119 standard; cDNA; 3213 BP.

XX AC AA42119;

XX AC AA42119; (first entry)

XX 31-JAN-1997 (first entry)

XX M-lats gene encoding large tumour suppressor.

XX Mouse; m-lats gene; large tumour suppressor; fetal brain;
XX protein-serine/threonine-kinase; cell proliferation; antisense;
XX dominant-negative; cancer; degenerative disorder; trauma;
XX growth deficiency; therapy; antitumour; vulnery; diagnostic;
XX Transgenic plant; transgenic animal; growth; senescence; ds.

XX Mus musculus.

XX OS
XX Key Location/Qualifiers
XX CDS 1..2889

XX FT /*tag= a
XX FT /product= m-lats protein

XX WO9630402-A1.

XX PD 03-OCT-1996.

XX PF 26-MAR-1996; 96WO-US04101.

XX PR 27-MAR-1995; 95US-0411111.

XX PA (UYUA) UNIV YALE.

XX PI Tao W, Wang W, Xu T, Yu W, Zhang S;

DR WPI: 1996-455275/45.
 DR P-PSDB; AA05179.
 XX New isolated large tumour suppressor gene - used to develop prods:
 PT for inhibiting cell proliferation or for enhancing proliferation
 PT
 XX
 PS Disclosure: Page 126-130; 215pp; English.
 XX
 CC This sequence encodes a mouse large tumour suppressor m-lats protein,
 CC and has been isolated from a newborn mouse brain phage lambda-ZAP
 CC cDNA library using a 2.2-kb DNA probe from the Drosophila lats gene
 CC (AA742117). A homologous mouse sequence has also been isolated
 CC (m-lats2, AA742120). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 SQ Sequence 3213 BP; 946 A; 762 C; 697 G; 808 T; 0 other;
 Query Match 99.9%; Score 3209.8; DB 17; Length 3213;
 Best Local Similarity 99.9%; Pred. No. 0;
 Matches 3211; Conservative 0; Mismatches 2; Indels 0; Gaps 0;

QY 1 GTGCAACATTCAATTAACGGAACAAAGCTGGAAGGTTCTTAAGAGTCTCTAGTTCCT 60
 DB 1 GTGCAACATTCAATTAACGGAACAAAGCTGGAAGGTTCTTAAGAGTCTCTAGTTCCT 60

QY 61 CAGACAGCGGCCCATCTCTAGAGAAATGTGTTTATGTTCTGAAAGCCCACTCA 120
 DB 61 CAGACAGCGGCCCATCTCTAGAGAAATGTGTTTATGTTCTGAAAGCCCACTCA 120

QY 121 CAGCGGATAGGAAGACCTCTGTGATCGGCAATGTCAGCATTTGCTCAAGCTAC 180
 DB 121 CAGCGGATAGGAAGACCTCTGTGATCGGCAATGTCAGCATTTGCTCAAGCTAC 180

QY 181 CCAAGCAATGGACAGAGTGAACCCCAACCACTTCAAGTATGTTTCTCTCT 240
 DB 181 CCAAGCAATGGACAGAGTGAACCCCAACCACTTCAAGTATGTTTCTCTCT 240

QY 241 CCACCACTCCGAGAGCCAGACCCCACTCCCGAGGACCACTCCCGCTCCCGCTCA 300
 DB 241 CCACCACTCCGAGAGCCAGACCCCACTCCCGAGGACCACTCCCGCTCCCGCTCA 300

QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGAGTACGTAATCTCCGA 360
 DB 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGAGTACGTAATCTCCGA 360

QY 361 ATCTCCCTGTTCCACTGGGGGTGCGAGAGGGTACCCCTCCACCACTCTTACCACT 420
 DB 361 ATCTCCCTGTTCCACTGGGGGTGCGAGAGGGTACCCCTCCACCACTCTTACCACT 420

QY 421 TCTCCCATGAATCCCTTAGCGGTCTCAGAGGCCATTAGTCTGTTCAGTTGGTAGA 480
 DB 421 TCTCCCATGAATCCCTTAGCGGTCTCAGAGGCCATTAGTCTGTTCAGTTGGTAGA 480

QY 481 CAACCCATCATGACAGTACTAGCAAAATTTAACTTTACACCGGGCCACCTGGAGTT 540
 DB 481 CAACCCATCATGACAGTACTAGCAAAATTTAACTTTACACCGGGCCACCTGGAGTT 540

QY 541 CAGAATGGTGTGCTAGTCTGATTTTATCGTCACCAAAATGTCCCACTGGTTCGTG 600
 DB 541 CAGAATGGTGTGCTAGTCTGATTTTATCGTCACCAAAATGTCCCACTGGTTCGTG 600

QY 601 ACTCGGAGCCACCACTCTCCTATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 660
 DB 601 ACTCGGAGCCACCACTCTCCTATATCTCTGACCCAGCTAATGGACAAAGCCCTCTGCT 660

QY 661 TTACAAACAGGGGCTTGTGCTGCCACCATTATCGCCCAATGGAAGCTTCCTCAGTCG 720
 DB 661 TTACAAACAGGGGCTTGTGCTGCCACCATTATCGCCCAATGGAAGCTTCCTCAGTCG 720

QY 721 ATGATGGTCCCAACAGCAAGCATCAACATGAGGCTTTATAATATTAATGTCCTCGGA 780
 DB 721 ATGATGGTCCCAACAGCAAGCATCAACATGAGGCTTTATAATATTAATGTCCTCGGA 780

QY 781 CTGCAACAGCGCTGCGCCAGCTGCTTCTGCTCTGCGCAGTCTATCCCAAGCGGTGG 840
 DB 781 CTGCAACAGCGCTGCGCCAGCTGCTTCTGCTCTGCGCAGTCTATCCCAAGCGGTGG 840

QY 841 CATGAAATTCCTACATGCAACCTTAACATACAGTGAAGTCAAAATCTTTTAATAACCCA 900
 DB 841 CATGAAATTCCTACATGCAACCTTAACATACAGTGAAGTCAAAATCTTTTAATAACCCA 900

QY 901 TTAGCAAGTAGAGCAAGTCACTCTGCTAATCTCAGCCTTCTGCCACTACAGTCACTGCC 960
 DB 901 TTAGCAAGTAGAGCAAGTCACTCTGCTAATCTCAGCCTTCTGCCACTACAGTCACTGCC 960

QY 961 ATCACACCGCTCTCTTATTAACAGCCCGTGAAGACATGCGCGTCTTGAACCAAGAGTG 1020
 DB 961 ATCACACCGCTCTCTTATTAACAGCCCGTGAAGACATGCGCGTCTTGAACCAAGAGTG 1020

QY 1021 CAGACTGCTTTAGCCCCCAACCCATCTTCTTGGATGCCACAGCCAGTTCAGACTGTTCCAG 1080
 DB 1021 CAGACTGCTTTAGCCCCCAACCCATCTTCTTGGATGCCACAGCCAGTTCAGACTGTTCCAG 1080

QY 1081 CCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCTATCCCACTGTTGCTGAA 1140
 DB 1081 CCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCTATCCCACTGTTGCTGAA 1140

QY 1141 GTCTCAAGTATCAAGGTCCACCGCTTATCCAAACATCTGCTACACCAACCCA 1200
 DB 1141 GTCTCAAGTATCAAGGTCCACCGCTTATCCAAACATCTGCTACACCAACCCA 1200

QY 1201 TCTGTCCCTCCATATGAGTCAAGTGAAGCCCTGCAAGATGAACAGCTAGCTTACCC 1260
 DB 1201 TCTGTCCCTCCATATGAGTCAAGTGAAGCCCTGCAAGATGAACAGCTAGCTTACCC 1260

QY 1261 AAGGAAGATGATAGTGAAGAGTGGGACAGTGTGACTCTGGGGATTAAGAAAGAA 1320
 DB 1261 AAGGAAGATGATAGTGAAGAGTGGGACAGTGTGACTCTGGGGATTAAGAAAGAA 1320

QY 1321 CAGATTACAACTTCACTTATCAGTGTTCGGAAGATGAAGAGATGAAGAGAGAG 1380
 DB 1321 CAGATTACAACTTCACTTATCAGTGTTCGGAAGATGAAGAGATGAAGAGAGAG 1380

QY 1381 TCTCGGATTTCAGATTACTTCCCAAGGCTTTAAGTCTTCTATGAGCAGCAGCTAGAG 1440
 DB 1381 TCTCGGATTTCAGATTACTTCCCAAGGCTTTAAGTCTTCTATGAGCAGCAGCTAGAG 1440

QY 1441 AAGCTCCTGAAGTCTCATCAGAGCTGTGATCGGAAGAGCAGCTAGAAATGAATG 1500
 DB 1441 AAGCTCCTGAAGTCTCATCAGAGCTGTGATCGGAAGAGCAGCTAGAAATGAATG 1500

QY 1501 ATGCGGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAGAAA 1560
 DB 1501 ATGCGGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAGAAA 1560

QY 1561 GAGTCTAACTATATCTGCTTTAAAGGGCTTAAATGACAAAGTCTATGTTGTAAGATA 1620
 DB 1561 GAGTCTAACTATATCTGCTTTAAAGGGCTTAAATGACAAAGTCTATGTTGTAAGATA 1620

QY 1621 AAGACATATGGAATAGAGCGTTTGGTGAAGTCTGTAGCAGAAAGTCCGATCTAA 1680
 DB 1621 AAGACATATGGAATAGAGCGTTTGGTGAAGTCTGTAGCAGAAAGTCCGATCTAA 1680

QY 1681 GCTTCTATGCAACAAGACTCTTCAAGAGAAAGAGCTTCTGCTCGAAATCAGGTGGCT 1740
 DB 1681 GCTTCTATGCAACAAGACTCTTCAAGAGAAAGAGCTTCTGCTCGAAATCAGGTGGCT 1740

QY 1741 CATGTGAAGCGGAGAGGGATATCTTAGCAGAAAGCCGACAAATGAGTGGGTGGTCCGCTG 1800

Db	2821	CACACTCGATGAGATGATCAACACACAGCTCCGATGGAACACACGAGATCTAGTG	2880
Qy	2881	TATGTTTAAATAACTAGGAGATCATTTGAAGAAATTTGCAAGAGCGCTCAAGTGCAGGGT	2940
Db	2881	TATGTTTAAATAACTAGGAGATCATTTGAAGAAATTTGCAAGAGCGCTCAAGTGCAGGGT	2940
Qy	2941	TTTTGAAGTTTGGAGAAAATATGCAAAATGTGCAGAGATTTGTGCTCTGTGTACAAATA	3000
Db	2941	TTTTGAAGTTTGGAGAAAATATGCAAAATGTGCAGAGATTTGTGCTCTGTGTACAAATA	3000
Qy	3001	TTTTTATTTCCCTAAGTTATGGGAATTTGTTTAAAAATGTTAAATTAATTCACCCCTTTTAA	3060
Db	3001	TTTTTATTTCCCTAAGTTATGGGAATTTGTTTAAAAATGTTAAATTAATTCACCCCTTTTAA	3060
Qy	3061	TTCCAGTAATTTAGAAAAAATTCGTTAAAGGAAGTAAATATGAACTGAGTATTATAGTC	3120
Db	3061	TTCCAGTAATTTAGAAAAAATTCGTTATAGGAAGAAGTAAATATGAACTGAGTATTATAGTC	3120
Qy	3121	AATTCCTTGGTACTTAAAGTACTTAAAAAGAGAAGCTGGTATCTTTTGTATATATAATAA	3180
Db	3121	AATTCCTTGGTACTTAAAGTACTTAAAAAGAGAAGCTGGTATCTTTTGTATATATAATAA	3180
Qy	3181	ATAATTTTAAATCCCAAAAAAATAAAAAA 3213	
Db	3181	ATAATTTTAAATCCCAAAAAAATAAAAAA 3213	

RESULT 3

AAAT42118	AAAT42118 standard; cDNA; 3984 BP.
XX	XX
XX	AC AAT42118;
XX	AC AAT42118;
DT	31-JAN-1997 (first entry)
XX	XX
DE	H-lats gene encoding large tumour suppressor.
XX	XX
KW	Human; h-lats gene; large tumour suppressor; fetal brain;
KW	protein-serine/threonine-kinase; cell proliferation; antisense;
KW	dominant-negative; cancer; degenerative disorder; trauma;
KW	growth deficiency; therapy; antitumour; vulnery; diagnostic;
KW	transgenic plant; transgenic animal; growth; senescence; ds.
XX	XX
OS	Homo sapiens.
XX	XX
FH	Key Location/Qualifiers
FT	CDS 231..3623
FT	/*tag= a
FT	/product= Lats protein
XX	XX
PN	WO9630402-AL.
XX	XX
XX	03-OCT-1996.
XX	XX
PF	26-MAR-1996; 96WO-US04101.
XX	XX
PR	27-MAR-1995; 95US-041111.
XX	XX
PA	(UYUA) UNIV YALE.
XX	XX
PI	Tao W, Wang W, Xu T, Yu W, Zhang S;
XX	XX
DR	WPI; 1996-455275/45.
DR	P-PSDB; AAW05178.
XX	XX
PT	New isolated large tumour suppressor gene - used to develop prods.
PT	for inhibiting cell proliferation or for enhancing proliferation
XX	XX
PS	Claim: 6; Page 118-123; 215pp; English.
XX	XX
CC	This sequence encodes a human large tumour suppressor h-lats protein,
CC	and has been isolated from a fetal human brain phage lambda-gt10.
CC	CCNA library using a 2.1-kb DNA probe from the mouse m-lats gene.

This sequence encodes a human large tumour suppressor h-lats protein, and has been isolated from a fetal human brain phage lambda-gt10 cDNA library using a 2.1-kb DNA probe from the mouse m-lats gene.

3732	GCTCTGTGTACAAATATTATTTTCCCTAAATTTATGGAAATCCCTTTTAAATGTTTAAATTT	37391
3046	ATTCCACCCCTTTTAATTCAGTAATTTAGAAAAAATTCGTATAAGGAAAGTAAATTTATGAA	3105
3792	ATTCCAGCCGTTTAAATCAGTATTAGAAAAAATTCGTATAAGGAAAGTAAATTTATGAA	3851
3106	CTGAGTATTATAGTCAATCTTTGGTACTTAAAGTACTTAAAGACAGAG	3154
3852	CTGAATATTATAGTCAGTCTCTTGGTACTTAAAGTACTTAAAGTAAAGTAG	3900
RESULT 4		
AAZ51505	AAZ51505 standard; DNA; 3984 BP.	
IIID	AAZ51505;	
XX AC	AAZ51505;	
XX AC	AAZ51505;	
DT	21-JUN-2000 (first entry)	
XX DE	Human Lats (large tumour suppressor) DNA.	
XX DE	Human Lats; large tumour suppressor; cytostatic; vulnerary;	
XX KW	cell overproliferation inhibitor; cdc2; cell cycle-dependent.kinase;	
XX KW	treatment; prevention; screening; cancer; skin; ovarian tumour;	
XX KW	soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;	
XX KW	LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;	
XX KW	dysplasia; degenerative disorder; growth deficiency; physical trauma;	
XX KW	hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.	
OS	Homo sapiens.	
XX	Key	
FFH	Location/Qualifiers	
CDS	231..3623	
ET	/*tag= a	
FTT	/product= "Lats protein"	
XX	W0200010602-A1.	
XX	02-MAR-2000.	
XX	18-AUG-1999; 99WO-US19068.	
XX	18-AUG-1998; 98US-0096996.	
XX	18-AUG-1998; 98US-0096997.	
XX	(UYUA) UNIV YALE.	
XX	Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;	
PPI	Turenchalk GS, Stewart RA;	
XX	WPI; 2000-246496/21.	
DR	P-PSDB; AAY70390.	
XX	Use of lats proteins, complexes of lats and cdc2 for treating cancer	
XX	that is refractory to treatment by standard chemotherapy and radiation	
XX	therapy, and disorders associated with aberrant levels of cdc2 activity	
XX	Claim 44; Fig 12; 134pp; English.	
PS	The present sequence is a DNA encoding human Lats (large tumour	
XX	suppressor) protein which is a cell overproliferation inhibitor and a	
XX	negative regulator of cell cycle-dependent kinase cdc2/cyclin A.	
XX	The present sequence is useful for treating cancer that is refractory	
XX	to standard chemotherapy or radiation therapy such as hyperplasia,	
XX	metaplasia, or dysplasia, and disorders associated with aberrant	
XX	levels of cdc2 activity. Conditions treated by promoting cdc2 function	
XX	include degenerative disorders, growth deficiencies, hypoproliferative	
XX	disorders, physical trauma, lesions, and wounds. An animal model	
XX	preferably a mouse, in which a lats gene has been disrupted by homologous	
XX	recombination, e.g. a lats knock-out mouse, is used for screening	
XX	compounds that can be used to treat or prevent cancer, particularly	
XX	skin cancer, soft tissue sarcomas and ovarian tumours, and disorders	

CC associated with pituitary dysfunction e.g. luteinizing hormone (LH).
CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
SO Sequence: 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;

Query Match : 77.3%; Score 2483; DB 21; Length 3984;
Best Local Similarity 87.4%; Pred. No. 0;
Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;

Qy	1	GTGCAACATTCAATTTACCGGAAACAAGCTGGAAGGTTCTAAAGAGTCTCTAGTTCCT	60
Db	732	GTGCGAACAATCAGTTTAAACCGCAACAGAGCTGGAAGGTTCTAAAGATCCTTAGTTCCT	791
Qy	61	CAGAGACACGGCCCATCTCTTAGGAGAAAATGTGGTTTATCGTTGAAAGCCCCAACTCA	120
Db	792	CAGAGCATGGCCGCCACTAGGAGAAAGTGTGGCTATCAITCTGAGAGTCCCAGCTCA	851
Qy	121	CAGCGGATGTAGGAAGACCTCTGTCTGGATCCGGCATTCGACATTTGCTCAAGCTCAC	180
Db	852	CAGACAGATGTAGGAAGACCTTTGTCTGGAATCTGGTATATATCAGCATTTGTTCCAGCTCAC	911
Qy	181	CAAGCAATTTGGACAGAGAGTGAACCCGCCACACACCTCAAGTTAGGAGTCTTACTCCT	240
Db	912	CCTAGCACGGACAGAGTGAACCCGCCACACCTCAAGTAAGGAGTGTACTCCT	971
Qy	241	CCACCACCTCCGAGAGGCCAGACCCGACCTCCCGAGGACACACTCCCTCCGCCCTCA	300
Db	972	CCACCACCTCCGAGAGGCCAGACTCCCGCTCCGAGAGGTACAACTCCACTCCCTCCCTCA	1031
Qy	301	TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCGGA	360
Db	1032	TGGGAACCAAACTCTCAACAAAGGCTATTCTGGAACATGGAATACGTAATCTCCGGA	1091
Qy	361	ATCTCCCTGTTTCCACCTGGGGCGTGGCAGGAGGGGTACCTCCACACCTCTTACCACCT	420
Db	1092	ATCTCTCTGTTCCACCTGGGGCATGGCAAGAGGGCTATCTCTCCACACCTCTTCAACACT	1151
Qy	421	TCTCCCATGAAATCCCTTACCTAGCCAGCTCAGAGGGCCATTAGTTCTGTTCCAGTTGGTAGA	480
Db	1152	TCCCCATGAAATCCCTTACCTAGCCAGCAGAGAGGCATTAGTTCTGTTCTGTTGGCAGA	1211
Qy	481	CAACCCATCATATGCAGAGTACTAGCAAAATTTAACTTTACACAGGGGACCTGGAGTT	540
Db	1212	CAACCAATCATATGCAGAGTCTAGCAAAATTTAACTTTCCATCAGGAGACCTGGAATG	1271
Qy	541	CAGATGGTGGTGGTCAGTCTGATTTTATCGTGGACACAAA---TGTGCCACCTGGTTCT	597
Db	1272	CAGATGGTACTGGCAAACTGATTTTCATGATACACCAAAATGTTGTCCTGCTGGCACT	1331
Qy	598	GTGACTCGGCAGCCACACCTCCATATCCTCTGACCCGAGCTAATGGCAAAAGCCCTCT	657
Db	1332	GTGAATCGGCAGCCACACCTCCATATCCTCTGACAGCAGCTAATGGCAAAAGCCCTCT	1391
Qy	658	GCTTTACAACACAGGGGCTTCTGCTGCTCCACATCATTCGCCAATGAAAACGTTCTCTCAG	717
Db	1392	GCTTTACAACACAGGGGATTTCTGCTGCTCTCTGCTCATATACAATGGAAGTATTCCTCAG	1451
Qy	718	TCGATGATGGTGGCCCAACAGGACAGTCAATACATGGAGCTTTTATAATATTAATGTCCT	777
Db	1452	TCTATGATGGTGGCAACAGAAATAGTCAATACATGGAACATATATAACATTAGTGTACCT	1511
Qy	778	GGACTGCAACAGCTGGCCCCAGTCTGCTGCTGCTGGCGCAGTCAATCCCAAGCGGT	837
Db	1512	GGACTGCAACAAATTTGGCCCTCAGTCATCTTCTGCTCCAGCCCGAGTCATCCCGAGCAGT	1571
Qy	838	GGGATGAAATTCCTACATGGCAACCTCAACATACCAGTGAGTCAAAATCTTTTAATAAC	897
Db	1572	GGGATGAAATTCCTACATGGCAACCTCAACATACCAGTGAGTCAAAATCTTTTAATAAC	1631
Qy	898	CAATTAGGAAGTAGAGCAAGTCACTCTGCTAATTTCTCAGCCTTTCTGCCACTACAGTCACT	957
Db	1632	CAATTAGGAATAGAGCAAGTCACTCTGCTAATTTCTCAGCCTTTCTGCCACTACAGTCACT	1691

QY	958	GCATACACACCGCTCCTATTTCACAGCCCGTGAAGAAGCATCGCGCTCTTGAACACGAG	1017
Db	1592	GCAATTACACGACTCCTATTTCACAGCCGTGAAGAAGTATGCGTGTATTAAAAACGAG	1751
QY	1018	CTGAGAGACTGTTTAGCCCAACACCCATCCTCTTTGGATGCCAGCCAGTTCAGACTGTT	1077
Db	1752	CTACAGACTGCTTTTAGCACCTACACACCCCTCTTTGGATACCAAGGCAATTCAACTGTT	1811
QY	1078	CAGCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTTCATCCCACTGTTGCT	1137
Db	1812	CAACCCAGTCTCTTTCTGAGGAAACCGTTCAAATGTGACTGTGATGCCACCTGTTGCT	1871
QY	1138	GAAGCTCAAGCTATCAAGGTCCACCCGCTTATCCAAAACATCTGTACACCAAAAC	1197
Db	1872	GAAGCTCAAACTATCAAGGACCAACCAACCCCTACCCAAAACATCTGTGCACCAAAAC	1931
QY	1198	CCATCTGCTCCTCCATATGATGAGTCAAGTAAAGCCCTGCAAAAGATGAACACCTAGCTTA	1257
Db	1932	CCATCTGTCTCCTCCATAGAGTCAATCAGTAAGCCTAGCAAAAGAGATCAGCCAAAGCTTG	1991
QY	1258	CCCAAGGAAGATGATAGTGAGAAGAGTGCAGACAGTGGTGACTCTGGGGATAAAGAAAAG	1317
Db	1992	CCCAAGGAAGATGAGAGTGAAGAGTATGAAAATGTGTAGTGGGGATAAAGAAAAG	2051
QY	1318	AAACAGATTACAACTTACCTTATCAGTTCGGAAGAACAGAAAGATGAAGAACGAAGA	1377
Db	2052	AAACAGATTACAACTTACCTTATCAGTTCGGAAGAACAGAAAGATGAAGAACGAAGA	2111
QY	1378	GAGTCTCGGATTACAGAGTTACTCCCAACAGACGCTTTAAGTTCTTCATGGAGACACGTA	1437
Db	2112	GAACTCGTATTCAAAGTTATTCTCCTCAAGCATTTAAATTTCTTATGGAGCAACATGTA	2171
QY	1438	GAGAAGCTCTGAAGTCTCATCAGCAGCGTCTGCATCGGAAGAGACAGCTAGAAAATGAA	1497
Db	2172	GAAGATGACTCAAACTCATCAGCAGCGTCTACATCGTAAAGAACAAATTTAGAGNATGA	2231
QY	1498	ATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAAGATGCTTTGCCAG	1557
Db	2232	ATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAAGATGCTTTGCCAA	2291
QY	1558	AAAGAGTCTAATATATATGCTTTAAAGGGCTAAATTTGGACAGCTCATGTTCTGTAAG	1617
Db	2292	AAAGAACTCTAATATATATGCTTTAAAGGGCTAAATTTGGACAGCTCATGTTCTGTAAG	2351
QY	1618	ATPAAGACATTAGGAATAGGACGTTTGGTGAAGTCTGTCTAGCAAGAAAAGTTCGATACT	1677
Db	2352	ATPAAGACATTAGGAATAGGACGTTTGGTGAAGTCTGTCTAGCAAGAAAAGTTCGATACT	2411
QY	1678	AAAGCTTTCTATGCAACAAGACTCTTCSAAGAAAGAGCTTCTGCTCCGGAATCAGGTG	1737
Db	2412	AAAGCTTTCTATGCAACAAGACTCTTCSAAGAAAGAGCTTCTGCTCCGGAATCAGGTG	2471
QY	1738	GCTCATGTCAAGCGGAGGAGATATCCTTAGCAGAGCCGACAATGAGTGGGTGGTCCGC	1797
Db	2472	GCTCATGTCAAGCGGAGGAGATATCCTTAGCAGAGCCGACAATGAGTGGGTGGTCCGC	2531
QY	1798	CTGTACTACTCTTTCCAGGACAGGACAACTTGTACTTTGTGATGGACTACATCTCTGG	1857
Db	2532	CTATATTTATTTATCCCAAGATAAGGACAAATTTATCTTTGTAATGGACTACATCTCTGG	2591
QY	1858	GGGGATATGATGAGCTTATTAATTAGAATGGGATCTTTCCCTGAAAATCTGCCACGATTC	1917
Db	2592	GGGTATATGATGAGCTTATTAATTAGAATGGGATCTTTCCAGAAAAGTCTGGCACGATTC	2651
QY	1918	TACATAGCAAACTTACCTGTGCAGTTGAAAGTGTTCATAAATGGGTTTATTTCATAGA	1977
Db	2652	TACATAGCAAACTTACCTGTGCAGTTGAAAGTGTTCATAAATGGGTTTATTTCATAGA	2711
QY	1978	CATATTAACCTCATAACTTTTATTGATGCGGTGATGGCCATATTAATTAATGACTGACTTT	2037
Db	2712	GATATTAACCTCATAACTTTTATTGATGCGGTGATGGCCATATTAATTAATGACTGACTTT	2771
QY	2038	GGCTGTGTGACTCGGCTTCAGATGGACACATGACTCCAGTACTACAGAGTGGGATCAC	2097

[illegible]

3852 CTGAATATTATAGTCAGTCTTGCTACTTTAAAGTACTTAAAAAAGTAG 3900

RESULT 5
AAX87396
ID AAX87396 standard; cDNA; 7382 BP.
XX AC
AC AAX87396;
XX DT 08-OCT-1999 (first entry)
XX DE Human WART1 CDNA.
XX KW WART1; hWART1; WART orthologue; human; signal transduction;
KW protein kinase; cancer; tumour; diagnosis; therapy; ss.
XX OS Homo sapiens.
XX FH Key Location/Qualifiers
FH 5'UTR 1..399
FT FT /*tag= a
FT FT repeat_region 12..63
FT FT /*tag= b
FT FT /note= "contains 10 copies of GGC repeat, similar
FT FT to repeats that undergo expansion in human
FT FT diseases associated with neuronal
FT FT phenotypes"
CDS 395..3787
FT FT /*tag= c
FT FT 3'UTR 3788..7382
FT FT /*tag= d
FT FT repeat_region 6058..6346
FT FT /*tag= e
FT FT /note= "Alu-J subfamily repeat motif"
FT FT polyA_signal 7339..7344
FT FT /*tag= f
FT FT variation replace(978,6)
FT FT /*tag= g
FT FT /note= "results in Ala/Gly change"
FT FT variation replace(1840)
FT FT /*tag= h
FT FT /note= "silent polymorphism"
FT FT variation replace(3252..3253)
FT FT /*tag= i
FT FT /note= "deletion of 2 adenines, causes
FT FT frameshift"
XX XX
PN WO9937787-A2.
XX XX
PD 29-JUL-1999.
XX PD
PF 20-JAN-1999; 99WO-US01145.
XX PR
PR 21-JAN-1998; 98US-0072023.
XX XX
PA (SUGEN-) SUGEN INC.
XX PL Flanagan P, Plowman GD;
XX PL
DR WPI; 1999-458698/38.
DR P-PSDB; AAY06526.
XX XX
PT New nucleic acid encoding human orthologs of Drosophila WART
PT proteins, used to identify specific modulators for treating cancer
PT or for diagnosis
XX XX
XX Claim 1; Page 120-121; 137pp; English.
XX CC This is the nucleotide sequence of a cdna clone coding for a human
CC orthologue, i.e. hWART1 (see AAY06526), of Drosophila non-receptor
CC serine/threonine kinase WART. hWART1 cdna was isolated from a
CC human bone marrow cdna library using PCR fragments of hWART1 as
CC probes. 3 Polymorphisms were identified in the hWART1 gene: (1) a

CC nucleotide 978, resulting in an Ala/Gly change; (2) at nucleotide
 CC 1840, silent; and (3) at nucleotides 3252-3253, comprising a
 CC deletion of 2 adenosines, resulting in a C-terminal truncation of
 CC hWART1 in the putative kinase domain. The latter frameshift
 CC mutation was observed in 2 independent clones from human bone
 CC marrow cDNA. Truncation of hWART1 could play a role in disease
 CC progression. hWART1 shows strong expression in cell lines from
 CC non-small cell lung cancer, ovarian tumours, central nervous system
 CC tumors, renal tumours and breast tumours, and may provide a target
 CC for oncology drug development. Nucleic acids encoding full-length
 CC hWART1 and hWART1 polypeptides lacking one or more of amino acid
 CC segments 12-45, 55-151, 236-377, 404-520, 555-559, 601-702, 691-998
 CC and 1011-1086, or lacking one or more of the N-terminal domain,
 CC catalytic domain, or C-terminal domain are claimed, as well as
 CC hWART2 sequences (AY87357), hWART1 and hWART2 polypeptides,
 CC antibodies, a method for identifying modulators of hWART function,
 CC and use of such modulator compounds to treat an abnormal condition
 CC involving hWART signal transduction, especially cancer. Probes for
 CC detection of hWART nucleic acids are also claimed.
 XX
 SQ Sequence 7382 BP; 2291 A; 1458 C; 1417 G; 2216 T; 0 other;

Query Match 77.2%; Score 2481.4; DB 20; Length 7382;
 Best Local Similarity 87.3%; Pred. No. 0;
 Matches 2768; Conservative 0; Mismatches 386; Indels 15; Gaps 4;

QY 1 GTCAACATTCATACCAACAAAGCTGGAAGGTTCTAAAGAGTCTAGTTCCT 60
 DB GTGAGCAATTCAGTTAACCCGAACAGAGCTGGAAGGTTCTAAAGAAATCCTTAGTTCCT 955
 QY 61 CAGAGACACGGCCCATCTCTAGGAGAAATGTGTGTTATCGTCTCTGAAAGCCCAACTCA 120
 DB 956 CAGAGGATGCGCCGCCCATCTAGCAGAAAGTGTGCGCTATCATCTTGAGAGTCCCACTCA 1015
 QY 121 CAGCGGATGTAGAGACCTCTGTCTGGATCGGGATTCGAGCATTCGTCAGCTCAC 180
 DB 1016 CAGACAGATGTAGAGACCTCTGTCTGGATCTGGTATATCATCAGCATTTGTTCAAGCTCAC 1075
 QY 181 CCAAGCAATGCAGACAGTGAACCCGCCACCACTCAAGTTAGGAGTGTACTCCT 240
 DB 1076 CTTAGCAACGGCAGAGGTGAACCCGCCACCACTCAAGTAAAGAGTGTACTCCT 1135
 QY 241 CCACCACTCCGAGAGGCCAGACCCCACTCCCGAGGCGACCACTCCCTCCCTCCCTCA 300
 DB 1136 CCACCACTCCGAGAGGCCAGACTCCCTCCCAAGAGGTACAACCTCCACCTCCCTCCCTCA 1195
 QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGAGTACGTAATCTCCCGA 360
 DB 1196 TGGGAACCAACTCTCAACAAAGCGTATCTCGGAACATGGATACGTAATCTCCCGA 1255
 QY 361 ATCTCCCTGTTCACCTGGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCACT 420
 DB 1256 ATCTCTCTGTCACCTGGGCGATGCAAGAGGGCTATCCTCCACCACTCTCAACACT 1315
 QY 421 TCTCCCATGAATCCCTTAGCCAGGCTCAGAGGCCATTAGTCTGTCTCCAGTGTGTA 480
 DB 1316 TCCCCCATGAATCTCTTAATCAAGGACAGAGAGGCATTAGTCTGTCTGTGTGGCAGA 1375
 QY 481 CAACCCATCATCTCAGAGTACTAGCAAAATTTAACTTTACACCGGCGCACTGGAGTT 540
 DB 1376 CAACCAATCATCTCAGAGTCTTAGCAAAATTTAACTTTCAACAGGAGACCTGGGATG 1435
 QY 541 CAGAAATGGTGGTGTGAGTCTGATTTTATCGTGACCAAAA ---TGTCCCACTGGTTCT 597
 DB 1436 CAGAAATGGTGGTGTGAGTCTGATTTTATCGTGACCAAAAATTTGTCCTGCTGGCACT 1495
 QY 598 GTGACTGGGAGCCACCACTCCATATCCCTGACCCAGCTATGACACAAAGCCCTCT 657
 DB 1496 GTGAATGGGAGCCACCACTCCATATCCCTGACAGCAGCTAATGGACAAAGCCCTCT 1555
 QY 658 GCTTTTAAACACAGGGGCTTCTGCTGCTCCCACTATTCGCAATGGAAAGCTTCCTCAG 717
 DB 1556 GCTTTTAAACACAGGGGCTGCTGCTGCTTCGTCATATACAAATGGAAGTATTCCTCAG 1615

QY 718 TCGATGATGCTGCCCAACAGGAACAGCTATAACATGGAGCTTTTATAATATGTCCT 777
 DB 1616 TCTATGATGCTGCCCAACAGGAATAGTATACATGGAAGTATATAACATGATGTCCT 1675
 QY 778 GGACTGCAACAGCCCTGGCCCCAGTGTCTTCTGTCCTGCGCAGTCACTCCCAAGCGGT 837
 DB 1676 GGACTGCAACAAATTTGGCCTCAGTCTATCTTCTGCTCCAGCCAGTCACTCCCGGACGT 1735
 QY 838 GGGCATGAATTCCTACATGCGCAACCTAACATACAGTCTGAGTCAAAATCTTTTATAAC 897
 DB 1736 GGGCATGAATTCCTACATGCGCAACCTAACATACAGTCTGAGTCAAAATCTTTTATAAC 1795
 QY 898 CCATTAGGAAGTAGAGCAAGTCACTCTGCTTAATCTTACGCTTCTGCCACTACAGTCACT 957
 DB 1796 CCATTAGGAAGTAGAGCAAGTCACTCTGCTTAATCTTACGCTTCTGCCACTACAGTCACT 1855
 QY 958 GGCATCACACCGCTCTCTTATTAACAGCCGCTGAAAGAGTATGCGCTCTGAACCCAGAG 1017
 DB 1856 GCAATTTACACCACTCTTATTAACAGCCGCTGAAAGTATGCGTGTATTAACCCAGAG 1915
 QY 1018 CTGCAGCTGCTTTAGCCCCCAACCCATCTCTTGGATGCCACAGCAGTCTCAGACTGTT 1077
 DB 1916 CTACAGACTGCTTTAGCACCTTACACACCTCTTGGATACCACAGCAATTCAACTGTT 1975
 QY 1078 CAGCCTTACCCCTTTTCTGAGGCTACAGTTCAGTGTGCTGTCTATCCACTGTTGCT 1137
 DB 1976 CAACCCAGCTCTTTCTTCTGAGGGAACCGCTTCAATGTGACTGTGATGCCACTGTTGCT 2035
 QY 1138 GAAGTCTCAAGCTATCAAGTCCACCCGCTTATCCAAACATCTGCTACACCAAAAC 1197
 DB 2036 GAAGTCTCAAACTATCAAGGACCCACCCCTACCCAAACATCTGCTGCCACCAAAAC 2095
 QY 1198 CCATCTGCTCCCTCATATGAGTCAAGTAAAGTCAAGCCCTGCAAGAGTCAACAGCTAGCTTA 1257
 DB 2096 CCATCTGCTCCCTCATATGAGTCAAGTCAAGTCAAGCCCTGCAAGAGTCAACAGCTAG 2155
 QY 1258 CCAAGGAAGATGATAGTGAAGAGTGGGACAGTGGTGTCTCTGGGATTAAGAAAG 1317
 DB 2156 CCAAGGAAGATGAGAGTGAAGAGTGGGATGAGTGAAGTGGTGTCTGAGTGGGATTAAGAAAG 2215
 QY 1318 AAACAGATTAACACTTCACTTCACTGTTTGGGAAAAACAAAGAAAGTGAAGAACGAAGA 1377
 DB 2216 AAACAGATTAACACTTCACTTCACTGTTTGGGAAAAACAAAGAAAGTGAAGAACGAAGA 2275
 QY 1378 GAGTCTCGGATTCAGAGTCTACTCCCAAGGCTTTAAGTCTTCTATGAGGAGCAGCTA 1437
 DB 2276 GAATCTCGTATTCAAAGTTATCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGA 2335
 QY 1438 GAGAACGCTCTGAGTCTCATCAGCAGCGCTCTCATCGGAAGAGCAGCTAGAAAAATGAA 1497
 DB 2336 GAAATGTACTCAAACTCTCATCAGCAGCGTCTACATCGTAAAAACAAATAGAGATGAA 2395
 QY 1498 ATGATCGGGTGGATTTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAG 1557
 DB 2396 ATGATCGGGTGGATTTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAA 2455
 QY 1558 AAAGAGTCTAACTATATCTCTTTAAAGGCTTAAATGCAAGTCTATGTTTGAAG 1617
 DB 2456 AAAGAACTAATTAACATCCGCTTTAAAGGGCTTAAATGCAAGTCTATGTTTGAAG 2515
 QY 1618 ATAAAGACATTAAGAAATAGGAGCGTTTGGTGAAGTCTGTAGCAAGAAAGTCTGATACT 1677
 DB 2516 ATAAAGACATTAAGAAATAGGAGCGTTTGGTGAAGTCTGTAGCAAGAAAGTCTGATACT 2575
 QY 1678 AAAGCTTTGTATGCACAAAGACTCTTTCGAAAGAAAGACGCTTCTGCTCCGAAATCAGGTG 1737
 DB 2576 AAGCTTTGTATGCACAAAGACTCTTTCGAAAGAAAGATGCTTCTTCTTCCGAAATCAGTC 2635
 QY 1738 GCTCATGTGAAGGGGAGAGGATATCTCTAGCAAGCCCAATGATGGTGGTGGTCCGC 1797
 DB 2636 GCTCATGTGAAGGGTGAAGAGATATCTCTGCTGAAGCTGACAATGAATGGTGGTGGTTCGT 2695

Qy	1798	CTGTACTACTCTTTCCAGGACAGGACAACACTTGTACTTTGTATGGACATACATTTCCTGGG	1855
Db	2696	CTATATTATTCAATTCCAAGATAAGGACAATTTATTACTTTGTAAATGGACATACATTTCCCTGGG	2755
Qy	1858	GGGATATGATGAGCCTATTAATTAAGATGGGCATCTTTCTCGAANAATCTGGCACGATTC	1917
Db	2756	GGTGATATGATGAGCCTATTAATTAAGATGGGCATCTTTCCAGAAAAGCTGGGCACGATTC	2815
Qy	1918	TACATAGCAGAACCTTACCCTGTGCAGTTGAAGAGTGTTCAAAAAATGGGTTTTATTTCATAGA	1977
Db	2816	TACATAGCAGAACCTTACCCTGCAGTTGAAGAGTGTTCAAAAAATGGGTTTTATTTCATAGA	2875
Qy	1978	GATATTAAACCTGATTAACATTTTGTGACCGTGATGCCCATATAAAATTGACTGACTTT	2037
Db	2876	GATATTAAACCTGATTAATTTTGTGATGATCGTAGTGCTCATATTAATTGACTGACTTT	2935
Qy	2038	GGCTTGTGACTGGCTTCAGATGGACATGACTCCAAGTACTACCAGAGTGGGGATCAC	2097
Db	2936	GGCTCTGCACTGGCTTCAGATGGACACACGATTCCTAAGTACTATCAGAGTGSTGACCAT	2995
Qy	2098	CCACGGCAGATAGCATGGATTTCAAGTAAACGAATGGGAGATCCCTCCAAATGTCCGGTGT	2157
Db	2996	CCACGGCAGATAGCATGGATTTCAAGTAAAGTGGGGGATCCCTCAAGCTGTCTGATGT	3055
Qy	2158	GGGACACAGCTGAAGCCATGGAGCGGAGAGCTGTCGCCAGCACACAGCATGTCTAGCC	2217
Db	3056	GGAGACAGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCAGCACAGCATGTCTAGCA	3115
Qy	2218	CATTCTCTGTTGGGACTCCCAATTTATTTSCACCTGAAGTGTCTACCTGCGAACAGGATAT	2277
Db	3116	CATTCTTGGTTGGGACTCCCAATTTATTTGCACCTGAAGTGTCTACCAACAGGATAC	3175
Qy	2278	ACACAGCTGTGTACGTGTGGAGTGTGGTGTATTCTTTCTGGAATGTTGGTGGGACAA	2337
Db	3176	ACACAGTGTGTGATTTGGTGGAGTGTGGTGTATTCTTTTGAATGTTGGTGGGACAA	3235
Qy	2338	CTCTCTTCTTTGGCACAAACCCCATTAGAAAACAAAATGAAGGTTATCATCTGCGCAAAC	2397
Db	3236	CTCTCTTCTTTGGCACAAACACCATTAGAAAACAAAATGAAGGTTATCACTGCGCAAAC	3295
Qy	2398	TCCTACACATCCCCTCAAGCTAAGCTGAGTCTGGAAGCCTCTGACCTCATATTATCAAA	2457
Db	3296	TCCTTTCACATTCACACCAAGCTAAACTAGTCTGGAAGCTTCTGATCTTATTTATTA	3355
Qy	2458	CTGTGTGAGGACCGAAGACCGCTCGGCAAGAACGGTGTCTGATGAGATAAAGGCTCAT	2517
Db	3356	CTTTGCGGAGCCGGAAGATCGCTTAGCCAGAGATGGTGTCTGATGAATATAAAGCTCAT	3415
Qy	2518	CCATTTTTTAAGACCATCGATTCTCTAGTGTATGACATGAGACAGCAGTCTGCTTCATACATC	2577
Db	3416	CCATTTTTTAACAATTTGACTCTCCAGTGACCTGAGACAGCAGTCTGCTTCATACATT	3475
Qy	2578	CCTAAATCAGCATCCAACACATACATCCAAATTTGACCCCTGTTGATCCTGATAAATTG	2637
Db	3476	CCTAAATCAGCATCCAACACATACATCCAAATTTGATCTGTTGATCTGCTGATAAATTG	3535
Qy	2638	TGGAGGATGGGACGAGGAGGAGAAAATACAGTGACACTCGAGCGGATGGTATAAAAAT	2697
Db	3536	TGGAGTGATGAACGAGGAGGAHAATGAATGACACTCTCAATGATGGTATAAAAAT	3595
Qy	2698	GGGAGACACCGGAGACGCTTTCTATGAGTTCACCTTTCCGAGGTTTTTTGATGACAAT	2757
Db	3596	GGAAAGCATCCTGAACATGCAATCTTATGAATTTACCTTCCGAAGGTTTTTTTGATGACAAT	3655
Qy	2758	GGCTACCCATATAATTATCCAAGCCTATTGAGTATGAATACATTCATTACACAGGCTCA	2817
Db	3656	GGCTACCCATATAATTATCCGAAGCCTATTGAATATGAATACATTAATTCACAAGGCTCA	3715
Qy	2818	GAACAACAGTCTGATGAAGATGATCAACACACAGCTCCGATGCAAAACAAACCCGAGATCTA	2877
Db	3716	GACACAGCTCGATGAGATGATCAAAACACAGGCTCAGAGATTTAAAAATCCGATCTA	3775
Qy	2878	GTGTATGTTTAATAAACTAGGAGATCATTTGTA --- - - - GAATTTTCCAAAGGCTCTGAAGTG	2933

[illegible]

RESULT 6

AAX32981
ID AAX32981 standard: DNA: 2442 BP.

XX AAX32981:

DT 30-JUN-1999 (first entry)

Human warts protein encoding DNA.

Human; warts protein; Drosophila; protein kinase; epithelial cell;
KW
cell regulation; cell proliferation; cell differentiation; tumour;
KW

AA
OS
Homo sapiens.

PN WO9915558-A1.

01-APR-1999.

AA
PF 24-AUG-1998; 98WO-JP03739.

PR 24-SEP-1997; 97JP-0258689.

PA (SUME) SUMITOMO ELECTRIC IND CO.

PI Kishimoto T; Naqamine Y, Nishiyama Y, Niwa S, Saya H;

DR WPI; 1999-244389/20.

DR P-PSDB; AAY06890.

PT Gene of vertebrates homologous to a *Drosophila* gene

PS Claim 7; Page 44-52; 68pp; Japanese:

CC The invention relates to a human warts protein, homologous to Drosophila
CC warts gene expression product. The warts expression product is a protein
CC kinase involved in the regulation of proliferation and differentiation of
CC epithelial cells and suppression of tumour formation. The gene, its
CC expression product and antibodies are useful in the study of the
CC mechanisms of tumour development. The present sequence represents a DNA
CC encoding a human warts protein.

Sequence 2442 BP; 777 A; 528 C; 491 G; 646 T; 0 other;

Query Match	60.6%;	Score 1946.8;	DB 20;	Length.2442;
Best Local Similarity	87.9%;	Pred. No. 0;		
Matches 2146;	Conservative	0;	Mismatches 292;	Indels 4;

QY : 331 TCTGGGAACATGGAGTACGTAATCTCCCGAAATCTCCCCCTGTTCCACCTGGGGCGTGGCAG 390

Db 1 TCTGGAAACATGGAAATACGTAATCTCCCGAATCTCTCCTGTCCCACCTGGGGCATGGCAA 60

391 GAGGGGTACCCCTCCACCACCTCTTACCACCTTCTCCCATGAATCCCCCTAGCCAGGCTCAG 450

Fri Jan 17 11:16:58 2003

1141	CTACATCTTAAAAACAATTAGAGAAATGAATGATCGGGTTGGATTATCTCAAGATGCC	1200
1528	CAGGATCAAAATGAGAAGATGCTTTCCGAGAAAGAGTCTAACTATATCTGCTTAAAGG	1587
1201	CAGGATCAAAATGAGAAGATGCTTTCCGAGAAAGAGTCTAACTATATCTGCTTAAAGG	1260
1588	GCTAAATGACAAAGTCTATGTTGTAAAGATAAAGACATTAAGGAATAGGAGCGTTGGT	1647
1261	GCTAAATGACAAAGTCTATGTTGTAAAGATAAAGACATTAAGGAATAGGAGCGTTGGT	1320
1648	GAAGTCTGCTAGCAGAAAGTGCATTAAGCTTTGTATGCAACAAAGACTCTTCCA	1707
1321	GAAGTCTGCTAGCAGAAAGTGCATTAAGCTTTGTATGCAACAAAGACTCTTCCA	1380
1708	AAGAAAGACGCTCTGCTCCGAAATCAGGTGGCTCATGTGAAAGCGGAGGATATCTTA	1767
1381	AAGAAAGACGCTCTGCTCCGAAATCAGGTGGCTCATGTGAAAGCGGAGGATATCTTA	1440
1768	GCAGAAAGCGGAGTGAAGTGGTGGTCCGCTGTACTACTCTTTCCAGGACAAAGCAAC	1827
1441	GCTGAAGCTGACAAATGAATGGTAGTTCTGCTATATATTATTCATTCGAAGATAAGCAAT	1500
1828	TTGTACTTTGTGATGGACTACATCTCTGGGGGGATATGATGAGCCTTATTAATTAGAA	1887
1501	TTATATCTTTGTAATGGACTACATCTCTGGGGGGATATGATGAGCCTTATTAATTAGAA	1560
1888	GCATCTTTCTGAAATCTGGCAGGATCTACATAGCAGAACTTACCTGTGCAGTTGAA	1947
1561	GCATCTTTCTGAAATCTGGCAGGATCTACATAGCAGAACTTACCTGTGCAGTTGAA	1620
1948	AGTGTTCATAAATGGTTTTATTCATAGAGATATTAACCTGATACATTTTCATTTGAT	2007
1621	AGTGTTCATAAATGGTTTTATTCATAGAGATATTAACCTGATACATTTTCATTTGAT	1680
2008	CGTGTGGCCATTAATAATTTGACTGACTTTGGCTTGTGCTGCTGCTGCTGCTGCTGCT	2067
1681	CGTGTGGCTCATATAATTTGACTGACTTTGGCTTGTGCTGCTGCTGCTGCTGCTGCT	1740
2068	GACTCCAAAGTACTACAGAGTGGGATCACCCAGGAGATAGCATGATTTTCAGTAAAC	2127
1741	GATTTAAAGTACTATCAGAGTGGTGGATCCAGCATCCAGGAGATAGCATGATTTTCAGTAA	1800
2128	GAATGGGAGATCTTCCAAATTTGGTGGGAGACAGACTGAAGCCACTGGAGCGGAGA	2187
1801	GAATGGGAGATCTTCCAAATTTGGTGGGAGACAGACTGAAGCCACTGGAGCGGAGA	1860
2188	GCTGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	2247
1861	GCTGCTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG	1920
2248	GCACCTGAAGTGTCTTCCAAATTTGGTGGGAGACAGACTGAAGCCACTGGAGCGGAGA	2307
1921	GCACCTGAAGTGTCTTCCAAATTTGGTGGGAGACAGACTGAAGCCACTGGAGCGGAGA	1980
2308	GTTATCTTTGTAATGTTGGTGGGAGACAGCTCTTCTTGGCAGCAAAACCCATTTAGAA	2367
1981	GTTATCTTTGTAATGTTGGTGGGAGACAGCTCTTCTTGGCAGCAAAACCCATTTAGAA	2040
2368	ACACAAATGAAGTGTCTTCCAAATTTGGTGGGAGACAGCTGAAGCCACTGGAGCGGAGA	2427
2041	ACACAAATGAAGTGTCTTCCAAATTTGGTGGGAGACAGCTGAAGCCACTGGAGCGGAGA	2100
2428	AGTCTGAGGCTCTGACCTCATTTATCAACTGTCGAGGAGCAGAGACCGGCTCGGC	2487
2101	AGTCTGAGGCTCTGACCTCATTTATTAAGCTTTGCGGAGGAGCCGAGAGTCTGCTAGGC	2160
2488	AGAAGGCTGTGATGAGATAAGGCTCATCTTCTTAAAGCCATTCGATTTCTCTAGT	2547
2161	AGAAGGCTGTGATGAGATAAGGCTCATCTTCTTAAAGCCATTCGATTTCTCTAGT	2220
2548	GATCTGAGACAGCAGTCTGCTTACATCTTAAATCAGCATCCAAACAGATACATCC	2607
2221	GACCTGAGACAGCAGTCTGCTTACATCTTAAATCAGCATCCAAACAGATACATCA	2280

61	GAGGCTATCTCCACCACTTTTCAACACTTCCCCCAATGAATCCTCTTAAAGGACAG	120
451	AGGCGCATAGTCTGCTGAGTGTAGACAAACCATCATCATGACAGAGTACTAGCAA	510
121	AGAGCATATAGTCTGCTGAGTGTAGACAAACCATCATCATGACAGAGTACTAGCAA	180
511	TTTAACTTTACACAGGGGAGCTGAGTGTAGATGGTGGTGGTGGTGGTGGTGGTGGTGGT	570
181	TTTAACTTTACACAGGGGAGCTGAGTGTAGATGGTGGTGGTGGTGGTGGTGGTGGTGGT	240
571	GTGACCAAAA---TGTCCTCACTGCTGCTGACCTGCGCAGCAGCAGCAGCAGCAGCAGCAG	627
241	ATACACCAAAATGTTGCTGCTGCTGACCTGCGCAGCAGCAGCAGCAGCAGCAGCAGCAG	300
628	CTGACCCAGCTAATGACAAAGCCCTCTGCTTTCAACAGGGGCTCTGCTGCTGCTCA	687
301	CTGACAGCAGTATGACAAAGCCCTCTGCTTTCAACAGGGGCTCTGCTGCTGCTCA	360
688	CCATCTGCGCATGGAACGCTTCTCAGTGTGATGATGGTGGGAGGATGCTGCTGCTGCT	747
361	TCGTATATCAAAAGGAGTATTCCTCAGTGTGATGATGGTGGGAGGATGCTGCTGCTGCT	420
748	AACATGAGCTTTTATAATATTAATGTCCTGGAGTGAACAGCCTGGCCGAGTCTGCT	807
421	RACATGGAATATATAACATTTAGTGTACCTGGAGTGAACAAATTTGGCTGAGTCACT	480
808	TCGTCTGCGCAGTATCCCAAGCGGTGGGATGATGATGGTGGGAGGATGCTGCTGCTGCT	867
481	TCGTCTGCGCAGTATCCCAAGCGGTGGGATGATGATGGTGGGAGGATGCTGCTGCTGCT	540
868	ATACAGTGTGAGTCAAAATCTTTTAAATACCCATTAGAAGTGAAGCAAGTCACTGCT	927
541	ATACAGTGTGAGTCAAAATCTTTTAAATACCCATTAGAAGTGAAGCAAGTCACTGCT	600
928	AATCTCAGCCTTCTGCGCATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	987
601	AATCTCAGCCTTCTGCGCATACAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	660
988	GTGAAAGCATCGGCTCTGAAACAGAGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCT	1047
661	GTGAAAGCATCGGCTCTGAAACAGAGCTGACAGTGTGCTGCTGCTGCTGCTGCTGCTGCT	720
1048	TCTTGGATCCAGCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	1107
721	TCTTGGATCCAGCAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	780
1108	TCAAGTGTGCTGCTATCCACCTGTTGCTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	1167
781	TCAAGTGTGCTGCTATCCACCTGTTGCTGAGCTTCAAGCTTCAAGCTTCAAGCTTCAAGCT	840
1168	CCCTTATCCAAACATCTGCTGACCAAAACCCATCTGCTCCCTCCATATGAGTCAAGT	1227
841	CCCTTATCCAAACATCTGCTGACCAAAACCCATCTGCTCCCTCCATATGAGTCAAGT	900
1228	AAGCCTGCAAAAGTGAAGAGTGTGCTTACCAAGAGATGATAGTGAAGAGTGGG	1287
901	AAGCCTGCAAAAGTGAAGAGTGTGCTTACCAAGAGATGATAGTGAAGAGTGGG	960
1288	GACAGTGTGCTGCTGGGATTAAGAAAGAAACAGATTAACACTTCACTTCACTGTT	1347
961	GAAATGTTGATGTTGGGATTAAGAAAGAAACAGATTAACACTTCACTTCACTGTT	1020
1348	CGGAAACAAAGAGTGAAGAGTGTGCTTCCGAGTGTGCTGAGTGTGCTTCACTTCCACAG	1407
1021	AGGAAACAAAGAGTGAAGAGTGTGCTTCCGAGTGTGCTGAGTGTGCTTCACTTCCACAG	1080
1408	GCCTTTAAGTCTTCAATGAGCAGCAGTGAAGAGTGTGCTGAGTGTGCTTCACTTCCACAG	1467
1081	GCATTTAAATCTTTATGAGCAACATGTAAGAAATGATGAGTGTGCTTCACTTCCACAG	1140
1468	CTGCACTGGAAGACAGCTAGAAATGAAATGATGAGTGTGCTTCACTTCCACAGTGGC	1527

QY	2656	GAGGAAATATCAGTACACTCT-GAGCGATGGTATAAATGGGAAGCACCOCGACGA	2714
DB	1261	GAAGAAATCAATGACACTCTCAATGGATGGTATAAATGGGAAGCACCCTGAACA	1320
QY	2715	CGCTTCTATGATTCACCTTCGGAGGTTTGTGATGACAATGGCTACCCATA	2768
DB	1321	TGCATTCATGAATTTACCTCCGAAGGTTTTTGTGATGACAATGGCTACCCATA	1374
RESULT 8			
ABAO8740	standard; cDNA; 1357 BP.		
XX	AC	ABA08740;	
XX	DT	11-JAN-2002 (first entry)	
XX	DE	Human large tumour suppressor 1 homologue cDNA, SEQ ID NO:516.	
XX	KW	Human; cytokine; cell proliferation; cell differentiation; growth factor;	
XX	KW	haematopoiesis regulation; tissue growth; immunomodulator; activin;	
XX	KW	inhibin; chemotaxis; chemokinesis; thrombolysis; oncogenesis;	
XX	KW	proliferation; metastasis; cancer; tumour; haematopoietic disorder;	
XX	KW	myeloid cell disorder; lymphoid cell disorder; asthma; arthritis;	
XX	KW	chronic inflammatory condition; proliferative retinopathy;	
XX	KW	atherosclerosis; coronary heart disease; arterial ischaemia;	
XX	KW	bone disorder; osteoporosis; vascular growth disorder;	
XX	KW	tissue regeneration; wound healing; infection; immune disorder;	
XX	KW	cell culture; drug screening; gene therapy; antiinflammatory;	
XX	KW	antiasthmatic; antiarthritic; haemostatic; antiarteriosclerotic;	
XX	KW	cytostatic; osteopathic; vasotropic; cardiant; virucide; antibacterial;	
XX	KW	antifungal; vulnerary; antitumor; ss.	
XX	OS	Homo sapiens.	
XX	PN	WO200157188-A2.	
XX	PD	09-AUG-2001.	
XX	PF	05-FEB-2001; 2001WO-US03800.	
XX	PR	03-FEB-2000; 2000US-0496914.	
XX	PR	27-APR-2000; 2000US-0560875.	
XX	PA	(HYSE-) HYSEQ INC.	
XX	PI	Tang YT, Liu C, Drmanac RT;	
XX	DR	WPI; 2001-457740/49.	
XX	DR	P-PSDB; ABB11496.	
XX	PT	Human proteins and DNA encoding sequences useful for preventing,	
XX	PT	treating or ameliorating a medical condition in a mammalian subject	
XX	PT	e.g. arthritis and cancer	
XX	PS	Claim 1; Page 554-555; 1963pp; English.	
XX	CC	Sequences ABB10981-ABB12330 represent 1350 novel human polypeptides, and	
XX	CC	sequences ABB10981-ABB12330 represent nucleic acids encoding them. The	
XX	CC	invention also relates to vectors and recombinant host cells comprising a	
XX	CC	nucleotide of the invention, methods of producing the novel polypeptides,	
XX	CC	antibodies against the polypeptides, methods of detecting the nucleotides	
XX	CC	or polypeptides in a sample, and methods of identifying compounds which	
XX	CC	bind to polypeptides of the invention. Although novel, many of the	
XX	CC	polypeptides of the invention have homology to known proteins, thereby	
XX	CC	giving an insight into their probable biological activities, and hence	
XX	CC	potential therapeutic applications. The polypeptides of the invention may	
XX	CC	have various activities, including cytokine, cell proliferation or cell	
XX	CC	differentiation activities; stem cell growth factor activity;	
XX	CC	haematopoiesis regulatory activity; tissue growth activity;	
XX	CC	immunomodulatory activity; activin- or inhibin-related activities;	
XX	CC	chemotactic or chemokinetic activities; haemostatic, thrombotic or	
XX	CC	thrombolytic activities; receptor or ligand activities; or may be	

involved in oncogenesis, cancer cell proliferation or metastasis. Depending on their biological activities, polypeptides and nucleotides of the invention are useful for preventing, treating or ameliorating medical conditions, e.g., by protein or gene therapy. Such conditions include cancers, haematopoietic disorders (e.g., myeloid or lymphoid cell disorders), chronic inflammatory conditions (e.g., asthma or arthritis), proliferative retinopathy, atherosclerosis, coronary heart disease, arterial ischaemia, bone disorders (e.g., osteoporosis), and abnormal vascular growth. Polypeptides involved with tissue regeneration and repair (or nucleic acids encoding them) may be used to promote wound healing (e.g., of burns, incisions and ulcers), while those with immunomodulatory activities may be used in the treatment of viral, bacterial and fungal infections in addition to immune disorders. Polypeptides with growth factor activity may be used in cell cultures to promote cell growth. For example, such polypeptides may be used to manipulate stem cells in culture to give rise to neuroepithelial cells that can be used to augment or replace cells damaged by illness, autoimmune disease or accidental damage. The polypeptides and nucleotides may also be used in the diagnosis of the above conditions, and in drug screening techniques. The present sequence represents a cDNA encoding a novel human polypeptide of the invention.

Sequence 1357 BP; 436 A; 254 C; 285 G; 382 T; 0 other;

Query Match 31.5%; Score 1012; DB 22; Length 1357;
Best Local Similarity 85.7%; Pred. No. 3.8e-242;
Matches 1164; Conservative 0; Mismatches 180; Indels 14; Gaps 3;

QY 1542 AAAGATGCTTGGCCAGAAAGAGCTAACTATATTCCTCTTAAAGGGCTAAATGGACAA 1601
DB 1 AAAGATGCTTGGCCA-AAAGAACTCTAATACATCCCTCTTAAAGGGCTAAATGGACAA 59
QY 1602 GTCTATGTTTGTAAAGATAAAGACATTAGCAATAGGAGCGTTTGGTGAAGTCGTCTAGC 1661
DB 60 GTCTATGTTTGTGAAGATAAAGACACTAGCAATAGGAGCAATTTGGTGAAGTCGTCTAGC 119
QY 1662 AGAAAGTCGATCTAAAGCTTTGTATGCAACAAGACACTCTTCTGAAAGAAAGACGTTCT 1721
DB 120 AAGAAAGTAGATACTAAGGCTTTGTATGCAACAACAACCTCTTCTGAAAGAAAGACGTTCT 179
QY 1722 GCTCCGAAATCAGTGCTCATGTGAAAGCGGAGAGGATATCTAGCAGAACCCGACAA 1781
DB 180 TCTTCGAAATCAAGTCGCTCATGTAAAGCTGAGAGAGATATCTGGCTGAAGCTGACAA 239
QY 1782 TGAGTGGGTGGTCCGCTCTACTCTTTCCAGGACAAAGACAACTTGACTTTGTGAT 1841
DB 240 TGAATGGGTAGTTCGTCTATATTTCATCCAGATAGGACAATTTATACITTTGTAAT 299
QY 1842 GGACTACATTCCTGGGGGATATGATGAGCCTATTATTAATTAGATGGCATCTTTCTGA 1901
DB 300 GGACTACATTCCTGGGGGATATGATGAGCCTATTATTAATTAGATGGCATCTTTCTGA 359
QY 1902 AAATCTGGCAGATTCTACATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAT 1961
DB 360 AAGCTGGCAGATTCTACATAGCAGAACTTACCTGTCAGTTGAAAGTGTTCATAAAT 419
QY 1962 GGGTTTATTCATAGAGATTTAAACCTGTAAACATTTTGAATGACCGTGATGCCATAT 2021
DB 420 GGGTTTATTCATAGAGATTTAAACCTGTAAACATTTTGAATGATCGTGTGTCATAT 479
QY 2022 TAAATTTGACTACATTTGGCTTTGTGCTCTGCTTCCAGATGGACACATGACTCCAAGTACTA 2081
DB 480 TAAATTTGACTACATTTGGCTTTGTGCTCTGCTTCCAGATGGACACATGACTCCAAGTACTA 539
QY 2082 CCAGATGGGATCACCCACCGGAGATAGTAGCATGGATTTTCAATACCAATGGGAGATCC 2141
DB 540 TCAGATGGTGGATTCACCGGAGATAGTAGCATGGATTTTCAATACCAATGGGAGATCC 599
QY 2142 TTCCAAATTTGGTGGGAGACACTCAAGCCTAGGAGCGGAGAGTGTCTGCCAGCA 2201
DB 600 CTCAGCTGTGATGTGGAGACAGACTCAAGCCATTAGAGCGGAGAGTGTGCCAGCA 659
QY 2202 CCAGCGATGTCTAGCCCAATTCCTGTTGGGAGCTCCCAATTTATTTGACACCTGAAGTCT 2261

[illegible][illegible]

Fri Jan 17 11:16:58 2003

Db 1465 CGATCAAGAGTTACTCCCTTATGCTTCAATTTCTTATGAGCAACACAGTGGAGAT 1524
Qy 1444 GTCTGAAGTCTCATCAGCAGCTGTGCTATCGAAGAACAGCTAGAAATGAATGATG 1503
Db 1525 GTATCAAAACCTACCAGCAGAGGTCAGCGGAGGCTACAGCTGAGCAGGAAATGGCC 1584
Qy 1504 CGGGTGGATATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTGGCCAGAAAG 1563
Db 1585 AAAGTGGGCTCTGTAGGCGGAGCAGCAGATGAGGAAGATCTCTACCAAGAGGAG 1644
Qy 1564 TCTAATATATATGCTTTAAAGGGCTTAAATGGACAAGTCTATGTTTGTAAAGATAAAG 1623
Db 1645 TCTAATATACACCGCTTGAAGGGCCAGATGGACAAGTCTTGTGAAATCAAG 1704
Qy 1624 ACATTAGGAATAGAGAGGTTTGGTGAAGTCTGTCTAGCAAGAAAGTGTGATTAAGCT 1683
Db 1705 ACTTAGCATCGTGCTTTGGGAAAGTGTGCTGCTGTGAAGCTGGACACTACGCT 1764
Qy 1684 TTGTATGACAAAGACTCTTGAAGAAAGAGCTTCTGCTCCGAATCAGGTGGCTCAT 1743
Db 1765 CTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTGTAACCGGAATCAAGTGGCCAT 1824
Qy 1744 GTAAAGCGGAGAGGATATCTTACGAGAACCGCACAAATGAGTGGTGTCCCGCTGTAC 1803
Db 1825 GTCAAGCTGAGAGGACATCTGCTGCTGAGCAGACAATGAGTGGTGTCAAACTTAC 1884
Qy 1804 TACTCTTTCCAGGACAAAGCAACTTGTGATGAGTACTACATCTCTGGGGGGAT 1863
Db 1885 TACTCTTTCCAGGACAAAGCAAGCTGTACTTTGTGATGAGTACTACATACCAAGCGGGAT 1944
Qy 1864 ATGATGACCTTAAATTAGAATGGGATCTTTCTTCAAAATCTGGCAGCATTTACATA 1923
Db 1945 ATGATGACCTTGTGATCAGATGAGGTCTTCTGAGCAGCTGGCCGCTTCTACAT 2004
Qy 1924 GCAGAACTTACCTGTGAGTGAAGTGTTCATAAATGGTGTTCATAGAGATAT 1983
Db 2005 GCAGAGTTGACCTGGCCATGAAAGTGTCCACAGATGGCTTTATCCACCGGAGATC 2064
Qy 1984 AAACCTGATACATTTTCAATGACCTGATGCGCATATTAATGACTGACTTTGGGTG 2043
Db 2065 AAGCTGACACATCTACTGACCTGATGATGATGATTAATGACTGACAGATTTTGGCTC 2124
Qy 2044 TGCATGCTTTCAGATGACATGACTTCCAGTACTACAGAGTGGGATCACCACGG 2103
Db 2125 TGCATGATTCAGGTGACTTCAAAATCCAAATCCAGAAAGGGAACACATGAGA 2184
Qy 2104 CAAGATACATGATTTTCAATCAAGTGGGAGATCTTCCAAATGTCGTGTGGGAC 2163
Db 2185 CAGGACAGCATGGACCGGCTGACCTCTGGGACGATGTTTCCAACTGCTGTGGAGAC 2244
Qy 2164 AGACTGAAGCCACTGAGCGGAGAGTCTGCTCCAGCAGCAGGATGCTAGCCCATCT 2223
Db 2245 AGTTAAGACCTTGAGCAGAGGCGGCGAAGAGCAGCAGGATGCTGTGCGACATCT 2304
Qy 2224 CTGGTGGACCTCCAAATTAATTCAGCTGAAGTGTCTACTGCGACAGATATACAG 2283
Db 2305 CTGTGCGGACACCAATTTACATCTCGGAGGTGCTTCCGCAAGGGTACACGAG 2364
Qy 2284 CTGTGACTGTGGAGTGTGCTGTATTTCTTGAATTTGTTGGTGGCAACCTCT 2343
Db 2365 CTCTGTGACTGTGGAGCTCGGTGTGATTTCTTTGAGATCTGTTGGGACCGCT 2424
Qy 2344 TTCTTGGCAAAACCCCATTAGAAACACAAATGAAGGTTATCATCTGGAACCTCTTA 2403
Db 2425 TTCTTGGCCCCCACCACACAGAGCGCAGCTGAAGGTGATCAACTGGGAGACCGGTG 2484
Qy 2404 CACATCTCTCAAGCTAAGTGTGCTGAGGCTCTGAGGCTCATTTATCAAACTGTGT 2463
Db 2485 CATATCTCTACGAGTGTGAGGCTTACCGCTGAGCGCCGAGACCTCATACAGAGCTGTG 2544
Qy 2464 CGAGGACCAAGACACCGCTCGGCAAGAACAGGTTGTGTGATGAATAAGGCTCATCCATTT 2523
Db 2545 TGCCTGCTGACTGCTCCCTGGCAGGGATGGGCGAGATGACCTCAAGGCACACCGCTTC 2604

Qy 2524 TTAAAGACCATCGATTTCTCTAGTGTATCTGAGACAGCAGTCTGTATACATCCCTAAA 2583
Db 2605 TTCAACACCATCGACTTTTCCGTTGACATCCGAAAGCAGGCTGCACCTAGCTCCCAACC 2664
Qy 2584 ATCAGCCATCCAAACAGATACATTCATTTCCACCTTGTGATCTGTATAATTTGTGGAGC 2643
Db 2665 ATCAGCCATCCCAATGGACACCTTCAATTTTGCACCGTGGATGAAGAAAGCCCTCGGAC 2724
Qy 2644 GATGCGAGGAGGAGAAATATCAGTGCACACTCTGAGCGGATGTATATAAATGGGAAG 2703
Db 2725 GAGGCCAGCGGAGAG---AGGCCAAGCGCTGGGACACGCTGGCCTCCGCCAGCAGCAAG 2781
Qy 2704 CACCCCGAGCAGCTTTCTATGATTCACCTTTCCGAGGTTTGTGATGACAATGGCTAC 2763
Db 2782 CATCCAGACAGCCTCTATGATTCACCTTCCGAGGTTCTTCGATGACAACGGCTAT 2841
Qy 2764 CCATATAATTAATCCAAAGCCTATTGATGATGATATACATT 2802
Db 2842 CCCTTCCGCTGCCAAGCCCTCAGAGCCCGCAGAGAGT 2880

RESULT 10

AAZ51507
ID AAZ51507 standard; DNA; 3155 BP.

XX AAZ51507;

XX 21-JUN-2000 (first entry)

XX Mouse Lats2 (large tumour suppressor) DNA.

XX Mouse; Lats2; large tumour suppressor; cytostatic; vulnary;
XX cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
XX treatment; prevention; screening; cancer; skin; ovarian tumour;
XX soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
XX LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
XX dysplasia; degenerative disorder; growth deficiency; physical trauma;
XX hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX Mus musculus.

XX Key Location/Qualifiers

XX CDS 1..2943

XX /tag= a "Lats2 protein"

XX /product= 970..1920

XX /tag= b

XX misc_feature 970..1920
FT /note= "This region is erroneously repeated in the mouse Lats2 DNA sequence shown in figure 14"

XX /tag= c

XX misc_feature 1921..2960
FT /note= "This region is missing in the mouse Lats2 DNA sequence shown in figure 14"

XX WO200010602-A1.

XX PN 02-MAR-2000.

XX PF 18-AUG-1999; 99WO-USI9068.

XX PR 18-AUG-1998; 98US-0096996.

XX PR 18-AUG-1998; 98US-0096997.

XX PA (UYUA) UNIV YALE.

XX Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

XX Turenchalk GS, Stewart RA;

XX WPI: 2000-246496/21.

XX P-PSDB; AAY70392.

XX Use of lats: proteins, complexes of lats and cdc2 for treating cancer

PT

that is refractory to treatment by standard chemotherapy and radiation therapy, and disorders associated with aberrant levels of cdc2 activity

Claim 44; Pages 112-117; 134pp; English.

The present sequence is a DNA encoding mouse Lats2 (large tumour suppressor) protein which is a cell overproliferation inhibitor and a negative regulator of cell cycle-dependent kinase cdc2/cyclin A. The present sequence is useful for treating cancer that is refractory to standard chemotherapy or radiation therapy such as hyperplasia, metaplasia, or dysplasia, and disorders associated with aberrant levels of cdc2 activity. Conditions treated by promoting cdc2 function include degenerative disorders, growth deficiencies, hypoproliferative disorders, physical trauma, lesions, and wounds. An animal model preferably a mouse, in which a lats gene has been disrupted by homologous recombination, e.g. a lats knock-out mouse, is used for screening compounds that can be used to treat or prevent cancer, particularly skin cancer, soft tissue sarcomas and ovarian tumours, and disorders associated with pituitary dysfunction e.g. luteinizing hormone (LH) hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

Sequence 3155' BP; 751 A; 924 C; 894 G; 586 T; 0 other;

```
Query Match      25.3%; Score 812.6; DB 21; Length 3155;
Best Local Similarity 65.1%; Pred. No. 3.3e-192;
Matches 1237; Conservative 0; Mismatches 644; Indels 18;
```

919	QY	CACTCTGCTTAATTCACAGCCTTCTGCCACTACAGTCACTGCCATCAGACCCGCTCCTATT	978
985	DB	CAACCTGAGCCCTCTACTGCCCGCCCCAACAGGCTACCGCGCTGAGCGGCCGACACATC	1044
979	QY	CAACAGCCCGTGAAAGAGATCGCCTCTCTGAAACCAGAGCTGCGAGCTGCTTTAGCCCA	1038
1045	DB	CTTCACCTCTGAGAGCGTGGTGTGCTGGGCCCGAGCCCCAGACGCGTGGGCC	1104
1039	QY	ACCACCTCTTGGATGCGACGCGATTCAGACTGTTTACGCTACCCCTTTTCTTGAG	1098
1105	DB	TGCGACCCCGCTGGGTGGCTGCGCCACAGACCTGCCACTGAGAGCCTGGAGACGAAG	1164
1099	QY	GGTACAGCTTCAAGTGTGCTGTCTATCCACCTGTTGCTGAAGCTCCAAGCTATCAAGGT	1158
1165	DB	GAGGCGCGCAGGCCACACCCCGTGGATGTGACTATGCGGCTCGAGCGCAGGTGC	1224
1159	QY	CCACCACCCCTTATCCAAACATCTGCTACACCAAAACCATCTGTCCCTCCATATGAG	1218
1225	DB	CCACGCCCTCCGTATCCAAAGCATTGCTGTGCCCAGTAAGTCTGAGCAGTACAGCGTG	1284
1219	QY	TCAGTAAGTAAGCCCTGCAAGATGAACAGCTAGCTTACCCAGGAAGATGATAGTGAG	1278
1285	DB	GACCTGGACAGCCTGTGCACCAGTGTGCAGCAGAGTCTGCGAGGGGCACCTGATCTAGAC	1344
1279	QY	AAGAGTGGCGGACAGT-----GGTGACTCTGGGGATAAGAAAAAGAAACAG	1323
1345	DB	GGAGTGAACAAGCCACAAAGGTGCGAAGGAGGACAAAGCTGCCAGACAAAAACGAG	1404
1324	QY	ATTACAACCTTCACTATCACTGTTGCGAAAAACAGAAAGATGAAGAACGAGAGAGTCT	1383
1405	DB	ATTACAGACCTCCCGTGCTCTCCGCAAGAATAGCAGATGAAGAGAAGAGAGAGTCT	1464
1384	QY	CGATTACAGAGTTACTCCCCACAGSCCTTAAAGTTCTTCATGAGCAGCAGCTAGACAAC	1443
1465	DB	CGATCAAGGTTACTCCCTTATGCCCTTCAATTTCTATGGAGCACACGCTGGAGAT	1524
1444	QY	GTCTGAACTCTATCAGCAGCGTCTGCTCGGAAGACGAGCTAGAAAAATGAAATGATG	1503
1525	DB	GTATCAAAACCTACACGACAAGGTACGCGGAGGCTACAGCTGCGAGCAGGAATGCC	1584
1504	QY	CGGTTGGATTATCTCAAGATCCCAGATCAATTGAGAAAGATGCTTTGCCAGAAAGAG	1563
1585	DB	AAGCTTGGGCTCTGTGAGGCGGAGCAGCAGATGAGGAAGATCCTCTCACCAGGAGG	1644
1564	QY	TCTAACTATTTCGCTTTAAAGGGCTAAAATGGACAAAGTCTATGTTTGTGAAGATAAG	1623

Db 2725 GAGCCAGCGGAGAG---AGCGCCAGGCGCTGGGACACGCTGGCCTCCCGCCAGCAGCAG 2781
QY 2704 CACCCGAGCAGCGCTTCTATGAGTTTCACTTTCGGGAGGTTTTTTTGTATGACATGGCTAC 2763
Db 2782 CATCCAGAGCAGCGCTTCTATGAGTTTCACTTTCGGGAGGTTTTTTTGTATGACATGGCTAT 2841
QY 2764 CCATATAATTATCCAAAGCCTATTGAGTATGAATACATT 2802.
Db 2842 CCCTTCGGTCCCGAGCCCTCAGAGCCGCGCAGAGAGT 2880

RESULT 11
ID AAX87397
XX AAX87397 standard; cDNA: 5276 BP.
AC AAX87397;
DT 08-OCT-1999 (first entry)
XX Human WART2 cDNA.

XX WART2; hWART2; WART orthologue; human; signal transduction;
KW protein kinase; cancer; tumour; diagnosis; therapy; ss.
XX Homo sapiens.

XX Location/Qualifiers
FH 1..374
FT 5'UTR /*tag= a
FT 375..3641
FT /*tag= b
FT 3642..5276
FT /*tag= c

XX W09937787-A2.
XX 29-JUL-1999.
XX 20-JAN-1999; 99WO-US01145.
XX 21-JAN-1998; 98US-0072023.
XX (SUGEN-) SUGEN INC.
XX Flanagan P, Plowman GD;
XX WPI; 1999-458698/38.
XX P-PSDB; AAY06527.

XX New nucleic acid encoding human orthologs of Drosophila WART
XX proteins, used to identify specific modulators for treating cancer
XX or for diagnosis

XX Claim 1; Page 122-123; 137pp; English.

XX This is the nucleotide sequence of a cDNA clone coding for a human
XX orthologue, i.e. hWART2 (see AAY06527), of Drosophila non-receptor
XX serine/threonine kinase WART. hWART2 cDNA was isolated from a
XX human bone marrow CDNA library using a PCR fragment of hWART1 as
XX probe. hWART2 is consistently expressed in human tumour cells
XX lines, except for most of the colon cancer lines examined. This
XX overexpression in tumour cells versus normal tissues may provide a
XX target for oncology drug development. Nucleic acids encoding
XX full-length hWART2 and hWART2 polypeptides lacking one or more of
XX amino acid segments 1-33, 43-139, 342-466, 467-480, 514-518,
XX and 974-1048, or lacking one or more of the N-terminal domain,
XX catalytic domain, or C-terminal domain are claimed, as well as
XX hWART1 sequences (AAX87396), hWART2 and hWART1 polypeptides,
XX antibodies, a method for identifying modulators of hWART function,
XX and use of such modulator compounds to treat an abnormal condition
XX involving hWART signal transduction, especially cancer. Probes for
XX detection of hWART nucleic acids are also claimed.

SQ Sequence 5276 BP; 1348 A; 1407 C; 1345 G; 1176 T; 0 other;
Query Match 24.5%; Score 787.2; DB 20; Length 5276;
Best Local Similarity 68.7%; Pred. No. 9.2e-186;
Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps 1;

QY 1220 CAGTAAAGTAAAGCCCTGCAAGATGAACAGCCTCGTGGCGGCCCAACGAGCGGCGGACAGAGCC 1279
Db 1987 CAGGATGAGCAGAGAGCTCGTGGCGGCCCAACGAGCGGCGGACAGAGCC 2046
QY 1280 AGAGTCCGAGCAGTGTGACTCTGGGATAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1339
Db 2047 GCAAAAGCGCCCAAGGGGACAAAGCGGAAAGGATAAAGAAAGGATTAAGAAAGGATTAAGAAAG 2106
QY 1340 TCACCTGTTTCGGAAAGAAAGAAAGATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1399
Db 2107 TTCCCGTCCGCAAAACAGCAG 2166
QY 1400 CCCACAGCCCTTTAAGTTCTTCATGGAGCAGCAGCAGTGTAGAGACGCTCTGAAGTCTCATC 1459
Db 2167 CGCATACGCCCTTTAAGTTCTTCATGGAGCAGCAGCAGTGTAGAGACGCTCTGAAGTCTCATC 2226
QY 1460 AGCAGCTGTGCATCGGAAGAAAGAAAGATGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 1519
Db 2227 AGCAGAGGTTAACCGGAGGCTGCAGCTGCAGCAAGAAATGCGCAAGAAATGCGCAAGAAATGCG 2286
QY 1520 AAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAGTCTACTATATATTCGTC 1579
Db 2287 AAGCTGAGCAGGAGCAGATCGGGAAGATCTCTACCAGAAAGAGTCTAATTAACAACAGGT 2346
QY 1580 TTAAGAGGCTAAATGGCAAGTCTATGTTTGAAGATTAAGAAAGATTAAGAAAGATTAAGAAAGAG 1639
Db 2347 TAAAGAGGCGCAAGATGGCAAGTCTATGTTTGAAGATTAAGAAAGATTAAGAAAGATTAAGAAAG 2406
QY 1640 CGTTTGTGAAGTCTGTCTAGCAAGAAAGTCTGATCTACTACTACTACTACTACTACTACTACTACT 1699
Db 2407 CCTTGGAGAGTGTGCTTGTGTAAGTGGACACTACGCGCTGTACGCCATGAAGA 2466
QY 1700 CTCTTCGAAAGAAAGAGTCTCTCGGAATCAGTGGCTCATGTGAAAGCGGAGAGG 1759
Db 2467 CCCTAAGGAAAGAGATGTCTTGAACCGGAATCAGTGGCGGCCACGTCAGGCGGAGAGG 2526
QY 1760 ATATCTAGCAGAGCGGACAAATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1819
Db 2527 ACATCTGGCGGAGGAGCAGACATGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 2586
QY 1820 AGGAAACTTGTACTTGTGATGGACTACATTCCTGGGGGGGATATGATGAGCCTATTAA 1879
Db 2587 AAGACAGCCTGTACTTGTGATGGACTACATTCCTGGGGGGGATATGATGAGCCTATTAA 2646
QY 1880 TTGAATGGGATCTTTCCTGAAATCTGGCAGATTTCTACATAGCAGAACTTACCTGTG 1939
Db 2647 TCCGGATGGAGGTCTTCCCTGAGCAGCTGGCCCGGTTCTACATCGCAGAGCTGACTTTGG 2706
QY 1940 CAGTTGAAGTGTTCATAAAATGGGTTTTTATCATAGAGATTAATAACCTGATTAACATTT 1999
Db 2707 CCATGTAGAGTGTCCACAAGATGGCTTCATCCAGCAGAGACATCAAGCCTGATTAACATTT 2766
QY 2000 TGATTGACCGTGTGCGCATATTAAATGACTGACTTTGGCTTGTGCACTGGCTTTCAGAT 2059
Db 2767 TGATAGATCTGGATGGTGCATTAATTAACCTACAGATTTCCGGCTCTGCACCTGGTTCAGT 2826
QY 2060 GGACACATGACTTCCAAAGTACTACAGAGTGGGGATCCCGCCAGGACAGATAGCATGATTT 2119
Db 2827 GGACTCAACAATTCAAAATATTACCAGAAAGGAGGAGCCATGTTCAGACAGGACAGCATGGAGC 2886
QY 2120 TCAGTAACGAATGGGGAGATCTCTCCATTTGCTGGTGTGGGACAGAGCTGAAGCCATGG 2179
Db 2887 CCAGCGAGCTCTGGGATGATGTGTCTAATCTGCGGTGTGGGACAGGCTGAAGACCCTAG 2946
QY 2180 AGCGGAGAGCTCTCGCCAGCAGCAGCAGCATGTCTAGCCCATTTCTCTGTTGGGACTCCCA 2239
Db 2947 AGCAGAGGCGCGGAGAGCAGCAGAGAGTGTCTGGCACTTCACTGCTGGTGGGAGCTCCAA 3006

QY 2240 ATTATATTGCACTGAAGTGTCTACTCGCAACAGGATATACACAGCTGTGTACTGGTGA 2299
 Db 3007 ACTACATCGCACCCGAGGTGCTCTCCGCAAGGGTACACTCACTGTGTACTGGTGA 3066
 QY 2300 GTGTTGGTGTATCTTCTGTAATGTGTGGTGGGACACCTCTTCTTGGCACAACCC 2359
 Db 3067 GTGTTGGTGTATCTTCTGTAATGTGTGGTGGGACACCTCTTCTTGGCACAACCC 3126
 QY 2360 CATTAGAAACAAATGAAGGTATCATCTGGAACACTTCTACACATCCCTCTCTCAAG 2419
 Db 3127 CCACAGAACCCAGCTGAAGGTGATCAACTGGGAGAACACGCTCCACATTCAGCCCAAG 3186
 QY 2420 CTAAGCTGAGTCTGAAGCTCTGACCTCATATATCAAACTGTGTGAGGACACAGAGACC 2479
 Db 3187 TGAAGCTGAGCTCTGAGCCAGGACCTCATCAACAGCTGTGTGCTCCGACAGACACC 3246
 QY 2480 GCCTCGGCAAGAACGGTGTGTAGATGAAGAGTCAATCAATTTTAAAGACCATCGATT 2539
 Db 3247 GCCTCGGCGGAATGGGCGGATGACCTGAGGCGGACCCCTTCTTCAAGCGCATGACT 3306
 QY 2540 TCTCTAGTCACTGAGACAGAGTCTGCTTATATATCCCTAAATCAGGATCCCAACAG 2599
 Db 3307 TCTCAGTGACATCCGAGGACGAGCCCTTCTGCTTCCACCATCAAGCCACCCATGG 3366
 QY 2600 ATACATCAATTTGACCTGTGTGATCTGATAAATTTGGAGCGATGCGCGAGGAGG 2659
 Db 3367 ACACCTCGAATTTGACCCCTGATAGTGAAGAGCCCTTGGACATGCCAGC---GAAG 3423
 QY 2660 AAATATCAGTGACACTCTGAGCGGATGGTATATAAATTTGGGAGCACCCTGAGCAGCGTT 2719
 Db 3424 GTAGCACCAGGCTTGGGACACACTCACCTCGCCCAATAACAAGCATCCTGAGCAGCAT 3483
 QY 2720 TCTATGAGTTCACCTTTTCGGAGTTTGTATGATCAATTTGCTACCATTAATATCAAA 2779
 Db 3484 TTTAGCAATTCACCTTCGGAAGTTTCTTGTATGATCAATTTGCTACCATTAATATCAAA 3543
 QY 2780 AGCTATTGAGTATGAATCATTTTATTCACAGGCTCTAGA 2819
 Db 3544 AGCTTCAGGACGACAGCTTCACAGGCTGAGAGCTAGA 3583

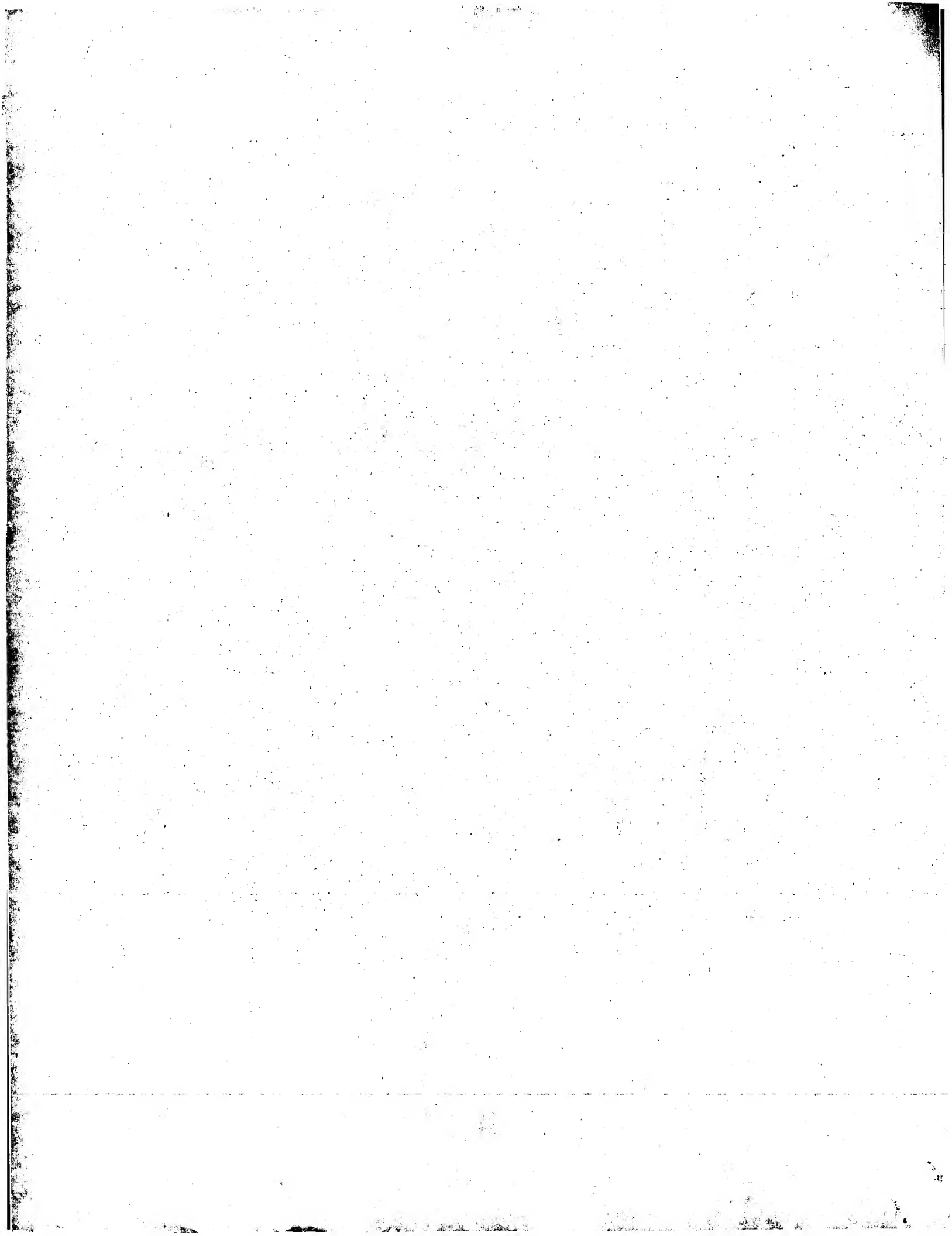
RESULT 12
 AAA59129
 ID AAA59129 standard; DNA; 5486 BP.
 XX
 AC AAA59129;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE DNA encoding a tumour suppressor protein hGHITS1.
 XX
 KW Human; growth hormone inhibited tumour suppressor protein; hGHITS;
 KW antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 KW diabetic nephropathy; cardiopathy; tumour; breast cancer;
 KW renal adenocarcinoma; colorectal cancer; leukaemia; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key
 CDS 387..3653
 FT
 FT /*tag= a
 FT /product= "tumour suppressor protein"
 XX
 PN EP1022333-A1
 XX
 PD 26-JUL-2000.
 XX
 PF 07-OCT-1999; 99EP-0119199.
 XX
 PR 25-JAN-1999; 99JP-0016223.
 XX
 PA (JCRP-) JCR PHARM CO LTD.

XX Koga J, Kono K, Zolotaryov FN;
 PI WPI: 2000-516013/47.
 DR P-PSDB: AAB07633.
 XX
 PT New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 PT diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 PT nephropathy or cardiopathy
 XX
 PS Claim 1: Page 18-28; 59pp; English.
 XX
 CC The present sequence encodes a human growth hormone inhibited tumour
 CC suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 CC antineoplastic pharmaceutical preparation. Probes for the hGHITS
 CC DNA sequences can be used in diagnostic pharmaceutical preparations.
 CC The diagnostic pharmaceutical preparations can be used for examining
 CC expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 CC angiopathy, diabetic nephropathy or cardiopathy, or in malignant
 CC tumours including breast cancer, renal adenocarcinoma, colorectal
 CC cancer, and leukaemia. Antibodies against the proteins can be used in
 CC a tumour suppressor gene.
 XX
 SQ Sequence 5486 BP; 1378 A; 1452 C; 1390 G; 1266 T; 0 other;
 Query Match 24.5%; Score 787.2; DB 21; Length 5486;
 Best Local Similarity 68.7%; Pred. No. 9.4e-186;
 Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps 1;
 QY 1220 CAGTAAGTAAAGCCGCAAGATGAACAGCTGTACCTTACCCAAAGGAGATGATAGTGA 1279
 Db 1999 CAGGATGAGGAGAGCCCTCGTGGGCGCCCAACAGCCGCGGCGCACAGAGCC 2058
 QY 1280 AGAGTGGGACAGTGTGACTCTGGGATAAAGAAAGAAACAGATTACACTTACACTA 1339
 Db 2059 GCAAAAGCGCCAAAGGGGACAAAGCGGAGGATAAAGAGCAGATTTCAGACCTCTCCG 2118
 QY 1340 TCACGTGTTGCGAAACAAAGAGATGAAGACGAGAGAGTCTCGGATTACAGATTACT 1399
 Db 2119 TTCCGCTCCGAAACAGAGAGAGAGAGAGAGAGTACGATCAAGAGCTACT 2178
 QY 1400 CCCACAGGCTTTAAGTTCTTCATGGAGCAGCAGTGAAGAACCTCTCTGAAGTCTATC 1459
 Db 2179 CGCATACGCTTTAAGTTCTTCATGGAGCAGCAGTGGAGATGTCATCAAAACCTACC 2238
 QY 1460 AGCAGCGTCTGCATCGGAGAGCAGCTAGAAAATGAATGATCGGGTGGATTATCTC 1519
 Db 2239 AGCAGAGGTTAAACCGGAGGCTGCAGCTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTG 2298
 QY 1520 AAGATCCCGCAGATCAATGAGAGATGCTTGGCAGAAAGAGTCTAACTATATTCGTC 1579
 Db 2299 AAGCTGAGCAGGAGCAGATGCGAAGATCCCTTACCAAGAAAGAGTCTAATTCACACAGT 2358
 QY 1580 TTAAGAGGCTAAAATGGACAGTCTATGTTTGAAGATAAAGACATAGGAATAGGAG 1639
 Db 2359 TAAAGAGGCGCAAGATGACAGTCTATGTTTGAAGATCAAAACCTGGGGATCGGTG 2418
 QY 1640 CGTTGGTGAAGTCTGTCTAGCAAGAAAGTCTGATCTAAAGCTTTGTATGCAACAAAGA 1699
 Db 2419 CTTTGGAGAAAGTGTGCTTAAAGTGGACACTCAGCGCCCTGACGCCATGAAGA 2478
 QY 1700 CTCCTCGAAAGAAAGACGTTCTGCTCCGAAATCAGTGGCTCATGTGAAAGCGAGAGGG 1759
 Db 2479 CCCTAAGGAAAGAGATGCTCCTGAACCGGAATCAGTGGCCCGCAGTCAAGCCGAGAGGG 2538
 QY 1760 ATATCTAGCAGAGCCGACAAATGAGTGGTGGTCCGCTGTACTACTCTTCCAGGAGA 1819
 Db 2539 ACATCTGCGCGAGCAGACAAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGAGA 2598
 QY 1820 AGGACAACTTGTACTTTGTGATGAGTACTATCTCTGGGGGATATGATGAGCCTATTAA 1879
 Db 2599 AAGACAGCCTGTACTTTGTGATGAGTACTATCTCTGGTGGGACATGATGAGCCTGTGA 2658

XX	DNA encoding a tumour suppressor protein hGHITS2.
DE	
XX	
KW	Human; growth hormone inhibited tumour suppressor protein; hGHITS;
KW	antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
KW	diabetic nephropathy; cardiopathy; tumour; breast cancer;
KW	renal adenocarcinoma; colorectal cancer; leukaemia; ss.
XX	
OS	Homo sapiens.
XX	
FH	Location/Qualifiers
FT	CDS 387..3653
FT	/**tag= a
FT	/product= "tumour suppressor protein"
XX	
PW	EPI022333-A1.
XX	
PD	26-JUL-2000.
XX	
PF	07-OCT-1999; 99EP-0119199.
XX	
PR	25-JAN-1999; 99JP-0016223.
XX	
PA	(JCRP-) JCR PHARM CO LTD.
XX	
PI	Koga J, Kono K, Zolotaryov FN;
XX	
DR	WPI: 2000-516013/47.
DR	P-FSDB; AAB07664.
XX	
PT	New human growth hormone inhibited tumour suppressor genes 1 and 2 for
PT	diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
PT	nephropathy or cardiopathy
XX	
PS	Claim 1; Page 28-39; 59pp; English.
XX	
CC	The present sequence encodes a human growth hormone inhibited tumour
CC	suppressor (hGHITS) protein. The hGHITS proteins can be used in an
CC	antineoplastic pharmaceutical preparation. Probes for the hGHITS
CC	DNA sequences can be used in diagnostic pharmaceutical preparations.
CC	The diagnostic pharmaceutical preparations can be used for examining
CC	expression of hGHITS genes in dwarfism, gigantism, acromegaly,
CC	angiopathy, diabetic nephropathy or cardiopathy, or in malignant
CC	tumours including breast cancer, renal adenocarcinoma, colorectal
CC	cancer, and leukaemia. Antibodies against the proteins can be used in
CC	a diagnostic pharmaceutical preparation for examining expression of
CC	a tumour suppressor gene.
XX	
SQ	Sequence 5486 BP; 1380 A; 1451 C; 1388 G; 1267 T; 0 other;
	Query Match 24.5%; Score 787.2; DB 21; Length 5486;
	Best Local Similarity .687%; Pred. No. 9.4e-186;
	Matches 1099; Conservative 0; Mismatches 498; Indels 3; Gaps
QY	1220 CAGTAAGTAAAGCCTCAAGAATGAACGCCCTAGCTTACCACAAGAGATGATAGTGAGA 1279
Db	1999 CAGGCATGAGCAGACGCTCGCTGGGGCCCCAACAGCCGCGGGCGGCACAAGAGCC 2058
QY	1280 AGAGTCGGACAGTGGTGACTCTGGGGATAAAGAAAACAGATTACAACTTCACCTA 1339
Db	2059 GCAAAAGCCCAAGGGGGACAAAGCGGAAGGATAAAACGACATTCAGACCTCTCCG 2118
QY	1340 TCACCTCTCGGAARACAGRAGATGAAGACGACAGAGATCTCGGATTCAGAGTTACT 1399
Db	2119 TTCCCCTCCGAARACAGCAGAGACGAAGAGAGAGATCACGATCAAGAGCTTACT 2178
QY	1400 CCCACAGCCCTTTAAGTTCTTCATGGACGACGCTAGAGAACTCTCTGAAGTCTCATC 1459
Db	2179 CGCCATACGCCCTTTAAGTTCTTCATGGACGACGCTGGAGATGTCTATCAAACCTACC 2238
QY	1460 ACCACGCTCTGCATCGGAAGACGAGCTAGAAATGAATGATCGGGTTGGATTAFTCT 1519
Db	2239 ACGAAGAGTTAAACCGAGGCTGCAGCTGGAGCAAGAAATGGCCAAGCTGCACTCTGTG 2298

1

2600	ATACATCAAAATTTGACCCCTGTTGATCTCGATAAATTTGTGGAGCGATGGCAGCGAGGAGG	2655
3379	ACACCTCGAAATTCGACCCCGTAGATGAAGAAAGCCCTTGGACGATGCCAGC---	3435
2660	AAAATATCAGTCACACTCTGAGCGGATGGTATAAAATGGGAAGACCCCGAGCAGCGCTT	2719
3436	GTAGCACCAGCCGTCGGACACACTCAGCTCGCCCAATAACAGCATCTCTGAGCAGCGCAT	3495
2720	TCATGAGTTACCTTTCCGAGGTTTTTTTGATGATCAATGGCTACCCCATATAATATCCAA	2779
3496	TTTACGATTCACCTTCGAGGTTCTTTGATGACATGGCTACCCCTTCGATGCCAA	3555
2780	AGCCATTGAGTATGATCAATTCATTACACAGGCTCAGA	2819
3556	AGCCTTCAGGAGCAGAAGCTTCACAGGCTGAGAGCTCAGA	3595
RESULT 14		
ID	AAZ61160	
XX	AAZ61160 standard; cDNA; 1961 BP.	
XX	AAZ61160;	
XX	30-MAY-2000 (first entry)	
XX	cDNA SS1771A encoding a partial protein kinase.	
DE	Kinase activity; molecular weight marker; isoelectric focusing marker;	
KW	peptide fragmentation control; cellular signal transduction; ss.	
XX	Homo sapiens.	
XX	Key	Location/Qualifiers
EH	FT	3..1769
FT	FT	/*tag= a
FT	FT	/note= "partial sequence"
XX	WO200008180-A2.	
PN	17-FEB-2000.	
PD	03-AUG-1999; 99WO-US17630.	
PF	04-AUG-1998; 98US-0095270.	
XX	11-SEP-1998; 98US-0099972.	
XX	(IMMV) IMMUNEX CORP.	
PA	Virca GD, Bird TA, Anderson DM, Marken JS;	
XX	WPI: 2000-195584/17.	
XX	P-PSDB; AAY69162.	
XX	New human kinase polypeptides and polynucleotides used as molecular	
PPT	weight markers and as controls for peptide fragmentation	
PPT	Claim 1; Page 9; 60pp; English.	
PS	The present sequence encodes a partial polypeptide which has kinase	
XX	activity. The kinase polynucleotides can be used to express the	
CC	polypeptides, and as probes to identify nucleic acids encoding	
CC	proteins having kinase activity. The kinase polypeptides and	
CC	fragmented polypeptides are used as molecular weight and isoelectric	
CC	focusing markers, and as controls for peptide fragmentation. They also	
CC	have a number of therapeutic uses as kinases play a central role in	
CC	cellular signal transduction. The polypeptides could also be used to	
CC	identify binding partner proteins. The polypeptides can also be used as	
CC	a reagent to identify any proteins that the polypeptide regulates, and	
CC	proteins with which it might interact. The polypeptides may also be	
CC	used for preparation of antibodies. The antibodies can be used in	
CC	assays to detect the presence of the protein, and to purify the protein	
XX	by immunoaffinity chromatography	



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:02 ; Search time 65.3715 Seconds
(without alignments)
15073.116 Million cell updates/sec

Title: US-09-763-334-3
Perfect score: 3213
Sequence: 1 gtcgaacattcaattacgc.....cccccaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2.6/ptodata/1/ina/5A_COMB.seq:*
- 2: /cgn2.6/ptodata/1/ina/5B_COMB.seq:*
- 3: /cgn2.6/ptodata/1/ina/6A_COMB.seq:*
- 4: /cgn2.6/ptodata/1/ina/6B_COMB.seq:*
- 5: /cgn2.6/ptodata/1/ina/PTOUS_COMB.seq:*
- 6: /cgn2.6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3213	100.0	3213	US-09-442-100-5	Sequence 5, Appli
2	2483	77.3	3984	US-09-442-100-3	Sequence 3, Appli
3	812.6	25.3	3155	US-09-442-100-7	Sequence 7, Appli
4	785.6	24.5	1961	US-09-509-902A-15	Sequence 15, Appli
5	713.8	22.2	1498	US-09-509-902A-6	Sequence 6, Appli
6	521.2	16.2	5720	US-09-442-100-1	Sequence 1, Appli
7	518.6	16.1	638	US-09-328-111-26	Sequence 26, Appli
8	376.6	11.7	678	US-09-328-111-66	Sequence 66, Appli
9	248.4	7.7	1935	US-08-878-989-11	Sequence 11, Appli
10	248.4	7.7	1935	US-09-272-796-11	Sequence 11, Appli
11	244.4	7.6	3018	US-08-860-150-6	Sequence 6, Appli
12	244.4	7.6	3018	US-09-338-132-6	Sequence 6, Appli
13	198.6	6.2	2101	US-08-860-150-1	Sequence 1, Appli
14	198.6	6.2	2101	US-09-338-132-1	Sequence 1, Appli
15	179.8	5.6	2160	US-09-588-256-1	Sequence 1, Appli
16	125.4	3.9	4739	US-08-685-871-1	Sequence 1, Appli
17	115.8	3.6	2311	US-08-712-709-6	Sequence 6, Appli
18	115.8	3.6	2311	US-09-111-444-6	Sequence 6, Appli
19	115.8	3.6	2311	US-09-541-228-6	Sequence 6, Appli
20	115.8	3.6	2370	US-09-031-295-1	Sequence 1, Appli
21	115.4	3.6	5053	US-08-685-576-2	Sequence 2, Appli
22	112.8	3.5	1599	US-09-256-465-1	Sequence 1, Appli
23	112.8	3.5	1599	US-09-167-322-3	Sequence 3, Appli
24	109	3.4	4363	US-08-685-576-5	Sequence 5, Appli
25	107	3.3	2549	US-09-467-082-3	Sequence 3, Appli
26	105	3.3	2726	US-08-422-699A-12	Sequence 12, Appli
27	105	3.3	2726	US-08-422-706B-12	Sequence 12, Appli

28	105	3.3	3182	1	US-08-484-044-11	Sequence 11, Appli
29	103.4	3.2	2511	2	US-08-422-699A-8	Sequence 8, Appli
30	103.4	3.2	2511	2	US-08-422-706B-8	Sequence 8, Appli
31	97.2	3.0	2706	2	US-08-630-822A-61	Sequence 61, Appli
32	97.2	3.0	2706	4	US-09-005-069-61	Sequence 61, Appli
33	97.2	3.0	2706	4	US-09-171-156A-20	Sequence 20, Appli
34	96.4	3.0	2104	4	US-09-313-930-1	Sequence 1, Appli
35	96.4	3.0	2196	1	US-08-313-274-1	Sequence 1, Appli
36	96	3.0	2244	3	US-09-094-714A-48	Sequence 48, Appli
37	96	3.0	2245	4	US-09-225-749-24	Sequence 24, Appli
38	88.2	2.7	2754	4	US-09-429-322-3	Sequence 3, Appli
39	78.8	2.5	1732	4	US-09-430-564-1	Sequence 1, Appli
40	78.2	2.4	1890	3	US-09-289-466-2	Sequence 2, Appli
41	78.2	2.4	1929	2	US-09-016-000-10	Sequence 10, Appli
42	77.6	2.4	2599	6	5266464-1	Patent No. 5266464
43	76.8	2.4	2557	4	US-08-464-954A-1	Sequence 1, Appli
44	76.6	2.4	1891	3	US-09-289-466-1	Sequence 1, Appli
45	73.6	2.3	2610	2	US-09-212-771-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-442-100-5
; Sequence 5, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Welyi
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATs
; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: PatentIn Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistrock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3213 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..2889

US-09-442-100-5

Query Match:		100.0%	Score 3213;	DB 4;	Length 3213;
Best Local Similarity:		100.0%	Pred. No. 0;		
Matches 3213;		Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;
Qy	1	GTGCAACATTC	CAATTAACCGA	AAACAAAGCT	GGAAGGTTCTA
Db	1	GTGCAACATTC	CAATTAACCGA	AAACAAAGCT	GGAAGGTTCTA
Qy	61	CAGAGACACG	CCCATCTCT	AGGAGAAATG	TGGTTTATCGT
Db	61	CAGAGACACG	CCCATCTCT	AGGAGAAATG	TGGTTTATCGT
Qy	121	CAGCGGATG	TAGGAAGAC	CTCTGTCTG	ATCCGGCATTC
Db	121	CAGCGGATG	TAGGAAGAC	CTCTGTCTG	ATCCGGCATTC
Qy	181	CCAGCAATG	AGAGAGAGT	GAACCCCTC	CCACACACCTC
Db	181	CCAGCAATG	AGAGAGAGT	GAACCCCTC	CCACACACCTC
Qy	241	CCACCACTC	CGAGAGG	CCAGACCTC	CCCGAGGACCA
Db	241	CCACCACTC	CGAGAGG	CCAGACCTC	CCCGAGGACCA
Qy	301	TGGGAACCA	AGCTCTC	AGAAAGCT	TAATCTCCGA
Db	301	TGGGAACCA	AGCTCTC	AGAAAGCT	TAATCTCCGA
Qy	361	ATCTCCCTG	TCCACCTG	GGGCTGG	AGAGGATAC
Db	361	ATCTCCCTG	TCCACCTG	GGGCTGG	AGAGGATAC
Qy	421	TCTCCCATG	ATCCCTC	AGGAGG	CTAGAGGCT
Db	421	TCTCCCATG	ATCCCTC	AGGAGG	CTAGAGGCT
Qy	481	CAACCATAT	CATGAGAG	TACTAGCA	AAATTAATCT
Db	481	CAACCATAT	CATGAGAG	TACTAGCA	AAATTAATCT
Qy	541	CAGATGGTG	GTGCTGCT	ATGCTG	CAACAAATG
Db	541	CAGATGGTG	GTGCTGCT	ATGCTG	CAACAAATG
Qy	601	ACTCGGAG	CCACCTC	CTGCTG	CTGCTGCTG
Db	601	ACTCGGAG	CCACCTC	CTGCTG	CTGCTGCTG
Qy	661	TTACAAAC	AGGGCTT	CTGCTG	CTGCTGCTG
Db	661	TTACAAAC	AGGGCTT	CTGCTG	CTGCTGCTG
Qy	721	ATGATGGT	GCCACAG	CAAGTAT	AAATGCTGGA
Db	721	ATGATGGT	GCCACAG	CAAGTAT	AAATGCTGGA
Qy	781	CTGCAAC	AGAGCTG	CTGCTG	CTGCTGCTG
Db	781	CTGCAAC	AGAGCTG	CTGCTG	CTGCTGCTG
Qy	841	CATGAAT	TCCTATG	CAACCTC	ATACAGT
Db	841	CATGAAT	TCCTATG	CAACCTC	ATACAGT
Qy	901	TTAGAAAG	TAGCAAG	CTGCTG	CTGCTGCTG
Db	901	TTAGAAAG	TAGCAAG	CTGCTG	CTGCTGCTG
Qy	961	ATCACAC	CCGCTCT	ATTAAC	AGCCGCTG
Db	961	ATCACAC	CCGCTCT	ATTAAC	AGCCGCTG

Qy	1021	CAGACTGCT	TTAGCCCC	CAACCCAT	CTTCTT
Db	1021	CAGACTGCT	TTAGCCCC	CAACCCAT	CTTCTT
Qy	1081	CCTACCCCT	TTTTCTT	GAGGTTAC	AGTTCAAG
Db	1081	CCTACCCCT	TTTTCTT	GAGGTTAC	AGTTCAAG
Qy	1141	GCTCCAA	GCTATCA	AGGTTCC	ACCCGCTT
Db	1141	GCTCCAA	GCTATCA	AGGTTCC	ACCCGCTT
Qy	1201	TCTGTCC	CTCATAT	GATGATC	AGTAAAG
Db	1201	TCTGTCC	CTCATAT	GATGATC	AGTAAAG
Qy	1261	AAGGAAG	ATGATG	AGAGAGT	GCGGACAG
Db	1261	AAGGAAG	ATGATG	AGAGAGT	GCGGACAG
Qy	1321	CAGATTAC	AACTTCA	CTTACCT	TATCCTT
Db	1321	CAGATTAC	AACTTCA	CTTACCT	TATCCTT
Qy	1381	TCTCGAT	TTCAGAG	TACTCCC	CAAGGCTT
Db	1381	TCTCGAT	TTCAGAG	TACTCCC	CAAGGCTT
Qy	1441	AACCTCT	CTGAAG	TCTCATC	AGAGGCTT
Db	1441	AACCTCT	CTGAAG	TCTCATC	AGAGGCTT
Qy	1501	ATCGGGT	TTGGAT	TATCTCA	AGATGCC
Db	1501	ATCGGGT	TTGGAT	TATCTCA	AGATGCC
Qy	1561	GAGTCTA	ACTATAT	TCTCTT	AAAGGCTT
Db	1561	GAGTCTA	ACTATAT	TCTCTT	AAAGGCTT
Qy	1621	AAGACAT	TAGGAAT	TAGAGG	TTGGTGA
Db	1621	AAGACAT	TAGGAAT	TAGAGG	TTGGTGA
Qy	1681	GCCTTCT	ATGCA	AAAGACT	CTTCCG
Db	1681	GCCTTCT	ATGCA	AAAGACT	CTTCCG
Qy	1741	CATGTGA	AGCGG	AGAGG	GATATCT
Db	1741	CATGTGA	AGCGG	AGAGG	GATATCT
Qy	1801	TACTACT	CTTCC	AGGAC	AACTTGT
Db	1801	TACTACT	CTTCC	AGGAC	AACTTGT
Qy	1861	GATATG	ATGAG	CTTAA	TAGATGG
Db	1861	GATATG	ATGAG	CTTAA	TAGATGG
Qy	1921	ATAGCAG	AACTT	ACATTTT	TCATG
Db	1921	ATAGCAG	AACTT	ACATTTT	TCATG
Qy	1981	ATTAAC	CTGAT	AACTT	TCATG
Db	1981	ATTAAC	CTGAT	AACTT	TCATG
Qy	2041	TTGTGCA	CTGCTT	CAGATG	GACACAT
Db	2041	TTGTGCA	CTGCTT	CAGATG	GACACAT

QY 2101 CGCAGATAGCATGATTTTCAGTACAGAAATGGGAGATCCTTCCAAATGTGCGGTGGG 2160
 Db 2101 CGCAGATAGCATGATTTTCAGTACAGAAATGGGAGATCCTTCCAAATGTGCGGTGGG 2160
 QY 2161 GACAGACTGAAGCCACTGAGCGGAGAGCTGCTGCCACACAGCGATGTCTAGCCCAT 2220
 Db 2161 GACAGACTGAAGCCACTGAGCGGAGAGCTGCTGCCACACAGCGATGTCTAGCCCAT 2220
 QY 2221 TCTCTGTTGGGACTCCCAATATATTGACCTGAAGTCTACTGCGAAGAGATATACA 2280
 Db 2221 TCTCTGTTGGGACTCCCAATATATTGACCTGAAGTCTACTGCGAAGAGATATACA 2280
 QY 2281 CAGCTGTGACTGGTGGAGTGTGGTGTATTTCTTTGTAATGTGTTGGGACAACT 2340
 Db 2281 CAGCTGTGACTGGTGGAGTGTGGTGTATTTCTTTGTAATGTGTTGGGACAACT 2340
 QY 2341 CTTTCTTGGCACAACCCCATAGAAACACAAATGAAGTTATCATCTGGCAAACTTCT 2400
 Db 2341 CTTTCTTGGCACAACCCCATAGAAACACAAATGAAGTTATCATCTGGCAAACTTCT 2400
 QY 2401 CTACACATCCCTCCTCAAGCTAAGTCTGCTGAAGCTCTGACCTCATTTCAAACTG 2460
 Db 2401 CTACACATCCCTCCTCAAGCTAAGTCTGCTGAAGCTCTGACCTCATTTCAAACTG 2460
 QY 2461 TGTGAGGACCAAGAACCGCTCGGCAAGACGGTGTGATGAGATAAGGCTCATCCA 2520
 Db 2461 TGTGAGGACCAAGAACCGCTCGGCAAGACGGTGTGATGAGATAAGGCTCATCCA 2520
 QY 2521 TTTTAAAGACCATCGATTTCTTAGTGTATCTGAGACAGCAGTCTGCTTCAATACCT 2580
 Db 2521 TTTTAAAGACCATCGATTTCTTAGTGTATCTGAGACAGCAGTCTGCTTCAATACCT 2580
 QY 2581 AAAATCACCATCCCAACAGATACATCCAAATTTGACCTCTGATCATATAATGTGG 2640
 Db 2581 AAAATCACCATCCCAACAGATACATCCAAATTTGACCTCTGATCATATAATGTGG 2640
 QY 2641 AGCGATGGCAGCGAGGAGGAAATATCAGTGACACTCTGAGCGGATGTTATATAATGG 2700
 Db 2641 AGCGATGGCAGCGAGGAGGAAATATCAGTGACACTCTGAGCGGATGTTATATAATGG 2700
 QY 2701 AAGCAGCCGAGCAGCCTTCTATGATGTCACCTTTCGAGGTTTTTGTGATGACATGGC 2760
 Db 2701 AAGCAGCCGAGCAGCCTTCTATGATGTCACCTTTCGAGGTTTTTGTGATGACATGGC 2760
 QY 2761 TACCATATAATTTATCAAGCTATTGATGATGATGATGATGATGATGATGATGATGAT 2820
 Db 2761 TACCATATAATTTATCAAGCTATTGATGATGATGATGATGATGATGATGATGATGAT 2820
 QY 2821 CAACAGTCTGATGAAGATGATCAACACACAAAGCTCCGATGGAACAAACCGAGATGATG 2880
 Db 2821 CAACAGTCTGATGAAGATGATCAACACACAAAGCTCCGATGGAACAAACCGAGATGATG 2880
 QY 2881 TATGTTTAAATAGGAGATCATTTGAAGATTTGCAAGAGGCTGAAGTGCAGGGGT 2940
 Db 2881 TATGTTTAAATAGGAGATCATTTGAAGATTTGCAAGAGGCTGAAGTGCAGGGGT 2940
 QY 2941 TTTTGAAGTTTGAAGAAATATGCAAAATGTCAGAGATTTCTGCTCTGTGTACAATA 3000
 Db 2941 TTTTGAAGTTTGAAGAAATATGCAAAATGTCAGAGATTTCTGCTCTGTGTACAATA 3000
 QY 3001 TTTTATTTTCCCTAAGTTATGGAAATTTGTTTAAATGTTAATTTATCCACCCCTTTAA 3060
 Db 3001 TTTTATTTTCCCTAAGTTATGGAAATTTGTTTAAATGTTAATTTATCCACCCCTTTAA 3060
 QY 3061 TTCAGTAAATTTAGAAAAATTTGTTATAGGAAGTAAATTTGACTGAGTATTATAGTC 3120
 Db 3061 TTCAGTAAATTTAGAAAAATTTGTTATAGGAAGTAAATTTGACTGAGTATTATAGTC 3120
 QY 3121 AATTTCTGCTACTTAAAGTACTTAAAGAGAGAGCTGCTGATCTTTTGTATATATAATA 3180
 Db 3121 AATTTCTGCTACTTAAAGTACTTAAAGAGAGAGCTGCTGATCTTTTGTATATATAATA 3180
 QY 3181 ATAATTTTAAATCCCAAAAAA 3213

Db 3181 ATAATTTTAAATCCCAAAAAA 3213
 RESULT 2
 US-09-442-100-3
 ; Sequence 3, Application US/09442100
 ; Patent No. 6359193
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Tian
 ; APPLICANT: Tao, Wufan
 ; APPLICANT: Wang, Welyi
 ; APPLICANT: Zhang, Sheng
 ; APPLICANT: Yu, Wan
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
 ; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent in Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/442,100
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/411,111
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mistock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 6523-003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEFAX: (212) 869-9741/8864
 ; TELEX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 3:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3984 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: cDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 231..3623
 ; US-09-442-100-3
 Query Match 77.3%; Score 2483; DB 4; Length 3984;
 Best Local Similarity 87.4%; Pred. No. 0;
 Matches 2769; Conservative 0; Mismatches 385; Indels 15; Gaps 4;
 QY 1 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGAGTTCTTAAAGAGTCTTAGTCTCT 60
 Db 732 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGAGTTCTTAAAGAGTCTTAGTCTCT 791
 QY 61 GAGACAGAGCGGCGGATCTCTAGGAGAAATGTGTTTATGTTCTTGAAGAGCGGCGGATCT 120
 Db 792 CAGAGGCGATGGCGGCGGATCTCTAGGAGAAATGTGTTTATGTTCTTGAAGAGCGGCGGATCT 851
 QY 121 CAGCGGAGTGTAGGAGAGCTCTGCTGATCCCGGCGGATGCGAGATTTGCTCAAGCTCAC 180
 Db 852 CAGACAGATGTAGGAGAGCTCTTGTGATCTGTGATCTATGATCTTGTGATCTTGTGATCT 911
 QY 181 CCAAGCAATGGACAGAGAGTGAACCCCGGCGGCGGATGCGAGATTTAGGAGTGTACTCTCT 240

Db 912 CCTAGCAACGAGACAGAGAGTGAACCCGCCACCACCACTCAAGTAAGGAGTGTTACTCCT 971
QY 241 CCACCACCTCCGAGAGGCGACGCCACCTCCCGGAGGACACACTCCCTCCCTCCCTCA 300
Db 972 CCACCACCTCCGAGAGGCGAGACTCCCTCCGAGGAGTACAACTCCCTCCCTCCCTCA 1031
QY 301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACGTAATCTCCGA 360
Db 1032 TGGGAACCAACACTCTCAACAAAGCGCTATTCTGGAACATGGAATACGTAATCTCCGA 1091
QY 361 ATCTCCCTGTTCCACTCGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCCT 420
Db 1092 ATCTCTCTGTTCCACTCGGCGCATGGCAAGAGGGGTATCTCTCCACCACTCTCAACT 1151
QY 421 TCTCCCATGAATCCCTTAGCAGGCTCAGAGGCGCTAGTTCGTTCCAGTTGGTAGA 480
Db 1152 TCCCCCATGANTCTCTAATCAGGACAGAGGGGTATGTTCTCTCTGTTGGCAGA 1211
QY 481 CAACCCATCATGACAGAGTACTAGCAAAATTTAACTTTACACAGGCGGACCTGGAGTT 540
Db 1212 CAACCAATCATGACAGAGTCTAGCAAAATTTAACTTTCCATCAGGAGACCTGGAATG 1271
QY 541 CAGAATGGTGGTTCAGTCTGATTTATCTGTCGACCAAAA---TGTCCCACTGGTCT 597
Db 1272 CAGAATGGTGGTTCAGTCTGATTTATCTGTCGACCAAAAATGTTGTCCTGCTGCACT 1331
QY 598 GTGACTCGGACGCCACCTCCATATCTCTGACCCAGCTTAATGGACAAAGCCCTCT 657
Db 1332 GTGAATCGGACGCCACCTCCATATCTCTGACCCAGCTTAATGGACAAAGCCCTCT 1391
QY 658 GCTTTACAAACAGGGGCTTCTGCTGCTCCACCATCATTCGCGCAATGGAAAGCTTCTCAG 717
Db 1392 GCTTTACAAACAGGGGCTTCTGCTGCTGCTCATATACAAATGGAAGTATTCCTCAG 1451
QY 718 TCGATGATGGTCCCAACAGGAACAGTATACATGAGGAGCTTAAATTAATGTCCT 777
Db 1452 TCTATGATGGTCCCAACAGGAATAGTATACATGGAAGTATTAATGATGAGCT 1511
QY 778 GAGCTGCAACAGCTGGCCCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 837
Db 1512 GAGCTGCAACAGCTGGCCCGAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1571
QY 838 GGGCATGAAATTCCTACATGCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 897
Db 1572 GGGCATGAAATTCCTACATGCACTCACTCACTCACTCACTCACTCACTCACTCACTCACT 1631
QY 898 CCATTAGGAAGTAGAGAGTCACTCTGCTTAATCTCAGCCCTTCTGCACTACAGTCACT 957
Db 1632 CCATTAGGAAGTAGAGAGTCACTCTGCTTAATCTCAGCCCTTCTGCACTACAGTCACT 1691
QY 958 GCCATCACCCGCTCTTATTCACAGCCCTGAAAGCATGCGGCTCTGAAACCCAGAG 1017
Db 1692 GCAATACACAGCTCTTATTCACAGCCCTGAAAGCATGCGGCTCTGAAACCCAGAG 1751
QY 1018 CTGACAGCTGCTTTAGCCCAACCCATCTCTTGGATGCAAGCCAGTTCAGCTGTT 1077
Db 1752 CTACAGACTGCTTTAGCCCAACCCATCTCTTGGATGCAAGCCAGTTCAGCTGTT 1811
QY 1078 CAGCCTACCCCTTTCTGAGGGTACAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCT 1137
Db 1812 CAACCCAGTCTTTCTGAGGAAACCGCTTCAATGTGACTGTGATGCACTGTTGCT 1871
QY 1138 GAAGCTCCAGCTATCAAGTTCACAGCCGCTTATCCAAACATCTGCTACACCAAAAC 1197
Db 1872 GAAGCTCCAAACTATCAAGGACCAACCAACCCCTTACCCAAACATCTGCTGCAACCAAC 1931
QY 1198 CCATCTCTCCCTTCCATGATGAGTAAAGTAAAGCCCTGCAAGATGAACAGCTAGCTTA 1257
Db 1932 CCATCTCTCTCCATGAGTCAATCAGTAAAGCCCTGCAAGATGAACAGCTAGCTTA 1991
QY 1258 CCNAGGAAGATGATGATGAGAGAGTGGGACAGTGGTGAATCTGGGGAATAAGAAAG 1317

Db 1992 CCCAAGGAGATCAGAGTGAAGAGTATGAAATGTTGATAGTGGGATAAAGAAAG 2051
QY 1318 AAACAGATTACAACCTTCACTCTTCCGAAAAACAAGAAAGATGAAGAACGAAAG 1377
Db 2052 AAACAGATTACAACCTTCACTCTTCCGAAAAACAAGAAAGATGAAGAACGAAAG 2111
QY 1378 GAGTCTCGATTACAGAGTACTCCCAAGAGGCTTAAAGTCTTCAAGGAGACAGCTA 1437
Db 2112 GAATCTCGATTACAAGTATCTCTCAAGCATTTAAATCTTTATGGAGCAACATGA 2171
QY 1438 GAGAACGCTTGAAGTCTCATCAGCAGCTCTGATCGGAAGAGAGCAGCTAGAAATGAA 1497
Db 2172 GAATCTCGATTACAAGTATCTCTCAAGCATTTAAATCTTTATGGAGCAACATGA 2231
QY 1498 ATGATGCGGGTGGAGTATCTCAAGATGCCAGAGTCAAAATGAGAAAGATGCTTTGCCAG 1557
Db 2232 ATGATGCGGGTGGAGTATCTCAAGATGCCAGAGTCAAAATGAGAAAGATGCTTTGCCAA 2291
QY 1558 AAAGAGTCTAACTATATTGCTTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 1617
Db 2292 AAAGAGTCTAACTATATTGCTTTAAAGGCTTAAAGGCTTAAAGGCTTAAAGGCTTAAAG 2351
QY 1618 ATAAAGACATTTAGGAATAGGAGGCTTTGGTGAAGTCTCTAGCAAGAAAGTGCATACT 1677
Db 2352 ATAAAGACATTTAGGAATAGGAGCATTTGGTGAAGTCTCTAGCAAGAAAGTGCATACT 2411
QY 1678 AAAGCTTTGATGCAACAAAGACTTTGAAAGAAAGAGCTTCTGCTCCGAAATFAGGTG 1737
Db 2412 AAGCTTTGATGCAACAAAGACTTTGAAAGAAAGAGCTTCTGCTCCGAAATFAGGTG 2471
QY 1738 GCTCATGTGAAGGCGAGAGGATATCTAGCAAGCCGACAAATGAGTGGTGGTCCGC 1797
Db 2472 GCTCATGTGAAGGCGAGAGGATATCTAGCAAGCCGACAAATGAGTGGTGGTCCGC 2531
QY 1798 CTGTACTACTCTTCCAGGACAAAGGACAACTTGTACTTTCTGATGAGTACATCTCTGGG 1857
Db 2532 CTATATTATTCATTCAGGATAGGACAAATTTATCTTCTTAATGAGTACATCTCTGGG 2591
QY 1858 GGGGATATGATGAGCTTAAATAGAAATGGGATCTTCTGAAAGTCTTCCGAAAGTCTGCAAGTTC 1917
Db 2592 GGTGATATGATGAGCTTAAATAGAAATGGGATCTTCTGAAAGTCTTCCGAAAGTCTGCAAGTTC 2651
QY 1918 TACATAGCAAGTCTTACCTGTCAGTGTAAAGTCTTCAAAATGGGTTTATTTATCATA 1977
Db 2652 TACATAGCAAGTCTTACCTGTCAGTGTAAAGTCTTCAAAATGGGTTTATTTATCATA 2711
QY 1978 GATATTAAACCTGATACATTTGATGACCTGATGGGCAATTAATTAATGAGTACTGCT 2037
Db 2712 GATATTAAACCTGATACATTTGATGACCTGATGGGCAATTAATTAATGAGTACTGCT 2771
QY 2038 GGCTTGTGCACTGGCTTCAGATGGACACATGACTCCCAAGTACTACCAAGTGGGATCAC 2097
Db 2772 GGCTTGTGCACTGGCTTCAGATGGACACATGACTCCCAAGTACTACCAAGTGGGATCAC 2831
QY 2098 CCAGGCAAGATAGATGATTTACATACGATGGGAGATCTTCCAAATTTCTCGGTG 2157
Db 2832 CCAGGCAAGATAGATGATTTACATACGATGGGAGATCTTCCAAATTTCTCGGTG 2891
QY 2158 GGGGACAGACTGAAGCCACTTGGAGCGGAGCTGCTCCGACAGCAGCAGGATGCTTAGCC 2217
Db 2892 GGGGACAGACTGAAGCCACTTGGAGCGGAGCTGCTCCGACAGCAGCAGGATGCTTAGCA 2951
QY 2218 CATCTCTGTTGGGACTCCCAATTAATTTGACCTGAAAGTCTGCTGCAAGAGGAT 2277
Db 2952 CATCTCTGTTGGGACTCCCAATTAATTTGACCTGAAAGTCTGCTGCAAGAGGAT 3011
QY 2278 ACACACTGCTGCTGAGTGGAGTGTGTTGTTATTTGTTGAAATGTTGTTGGGACAA 2337
Db 3012 ACACACTGCTGCTGAGTGGAGTGTGTTGTTATTTCTTTTGAATGTTGTTGGGACAA 3071
QY 2338 CTTCTCTTCTGGCAGAAACCCCATTAGAAACAAATGAGGTTATCATCTGCAAACT 2397
Db 3072 CTTCTCTTCTGGCAGAAACCCCATTAGAAACAAATGAGGTTATCATCTGCAAACT 3131

QY 1444 GTCTGAAGTCTCATCAGCAGCGCTGCTGCATCGGAAGAGCAGCTAGAAATGAATGATG 1503
 Db 1525 GTCTCAAAACCTACCAGCAGAGAGGTGAGCGGAGGCTACAGGTGAGCAGGAAATGGGC 1584
 QY 1504 CGGGTGGATATATCTCAAGATGCCAGGATCAAAATGAGAAGATGCTTTGCCAGAAAGAG 1563
 Db 1585 AAAGCTGGGCTCTGTGAGCGGAGCAGGAGCAGATGAGGAGATCTCTACAGAGAGGAG 1644
 QY 1564 TCTAATATATATCTGCTTTAAAGAGGCTTAAATGGACAAGTCTATGTTTGTAAAGATAAG 1623
 Db 1645 TCTAATATCAACCGGCTGAAGAGGCGCAAGATGGACAAGTCCATGTTGTGAATCAAG 1704
 QY 1624 ACATTAGGATAGAGCGTTTGGTGAAGTCTGTCTAGCAGAAAGATCGATATAAGCT 1683
 Db 1705 ACTTAGGATCGGCTTTGGGAGTGTGCTGCTGCTGTAAGCTGGACATCAAGCT 1764
 QY 1684 TTGTATGCAACAAAGACTCTTCGAAAGAAAGAGCTTCTGCTCGGAATCAGGTGGCTCAT 1743
 Db 1765 CTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTGAACCGGAATCAAGTGGCCAT 1824
 QY 1744 GTCAAGCGGAGAGGATATCTCTAGCAGAGCGGACAAATGAGTGGTGGTCCGCTGTAC 1803
 Db 1825 GTCAAGGCTGAGAGGACATCTGCTGGTGAAGCAGACAAATCAGTGGTGGTCAAACTCTAC 1884
 QY 1804 TACTCTTCCAGCAGACAGGACACTTGTACTTTGTGATGACTACATCTCTGGGGGGAT 1863
 Db 1885 TACTCTTCCAGCAGACAGGACACTTGTACTTTGTGATGACTACATCACCAGCGGGAT 1944
 QY 1864 ATGATGAGCTTAAATTAAGTGGGATCTTTCTGAAATCTGGCAGCATCTACATA 1923
 Db 1945 ATGATGAGCTGCTGATCAGGATGGAGTCTTCCCTGAGCAGCTGGCCCTCTTACATT 2004
 QY 1924 GCAGAACTTACCTGTGCTGAGTGAAGTGTTCATAAAATGGTGTTCATATAGATATT 1983
 Db 2005 GCAGAGTGTGACCTGGCCATTGAAAGTGTCCACAAGATGGCTTTATCCACCGGACATC 2064
 QY 1984 AAACCTGATAAACATTTTGTATGACGCTGATGGCCATATTAATGACTGACTTTGGCTG 2043
 Db 2065 AAGCTTGACAACTACTCTACGACCTGGATGCTATATTAAGCTGACAGATTTTGGCCTC 2124
 QY 2044 TGCACTGGCTTCAGATGGACATGACTCCCAAGTACTACAGAGTGGGATCACCCACGG 2103
 Db 2125 TGCACTGGATGAGTGGACTCACAATCCAGTACTACAGAAAGGACACATGAGA 2184
 QY 2104 CAAGATAGCATGATTCAGTACAGAAATGGGAGATCTTCCAAATGCTCGGTGGGGAC 2163
 Db 2185 CAGGACAGCATGAGCGCGGTGACCTCTGGGACGATGTTTCCCAACTGCTGCTGGAGAC 2244
 QY 2164 AGACTGAGCACTGGAGCGGAGAGCTGCTCGCCAGCAGCAGATGCTAGCCCATCT 2223
 Db 2245 AGTTAAAGACCTGGAGCAGAGGGCGCAGAAGCAGCACACAGAGTGCCTGGCAGATTCT 2304
 QY 2224 CTGGTGGGACTCCCAATATATTGACCTGGAAGTGTCTACTGCAACAGGATATACAG 2283
 Db 2305 CTGTGCGGACACAAATTAATCATCTCGGAGGCTGCTTCCGCAAGAGGTACAGCAG 2364
 QY 2284 CTGTGACTGGTGGAGTGTGTTATCTTTTGTGAATGTTGGTGGGCAACCTCTCT 2343
 Db 2365 CTCTGTGACTGGTGGAGCGTGGTGTGATCTCTTTGAGATGCTGTTGGGAGCGCCT 2424
 QY 2344 TTCTTGGCACAACCCCATAGAAACACAAATGAAGTTATCATCTGGCAAACTTCTCTA 2403
 Db 2425 TTCTTGGCCCCCCCCACAGAGAGCGCAGCTGAAGGTGATCAACTGGGAGACAGCGTG 2484
 QY 2404 CACATCTCTCTCAAGCTAAGCTGAGTCTCTGAAGCCTCTGACCTCATTAATCAAACTGTGT 2463
 Db 2485 CATATCTCTACGAGGTGAGGCTCAGGCTGAGCGCCGAGACCTCATCAGGAAGCTGTGC 2544
 QY 2464 CGAGGACCAAGAGACCGCTTCGCAAGACGGTCTGATGAGATAAAGGCTCATCCATTT 2523
 Db 2545 TGGCGGCTGACTGCGCGCTGGCAGGATGGGCGAGATGACCTCAAGGACACACCGCTTC 2604
 QY 2524 TTTAAGACCATCATTTCTCTAGTATCTGATCTGAGACAGCAGTCTGCTTCATACATCCCTAA 2583

Db 2605 TTCAACACCATTCAGCTTTTCCCGTGACATCCGAAGAGCAGCTCACCTAGTCCCAACC 2664
 QY 2584 ATCAGCGATCCACAGATACATCAATTTTCAGCCCTGTTGATCTCTGATAAATTTGGAGC 2643
 Db 2665 ATCAGCCACCCATGGACACCTCCAATTTTACCCGGTGTGATGAAGAAAGCCCTGGCAC 2724
 QY 2644 GATGGCAGCAGGAGGAGAAATATCAGTGACACTCTGAGCGGATGGTATATAAATGGGAG 2703
 Db 2725 GAGCGCAGCGGAGAG---AGCGCAAGCGCTGGGACACGCTGGCCTCCCCCAGCAGCAAG 2781
 QY 2704 CACCCGAGCAGCCTTCTATGATGATTCACCTTTCGAGGTTTTCGATGACATGGCTAC 2763
 Db 2782 CATCCAGAGCAGCGCTTCTATGATGATTCACCTTCCGAGGTTCTTCGATGACAAAGGCTAT 2841
 QY 2764 CCATATAATTTATCAAGCCTATTGAGTATGATATACATT 2802
 Db 2842 CCCTCCGCTGCCGAGCGCTCAGAGCCCGCAGAGAGT 2880

RESULT 4
 US-09-509-902A-15
 ; Sequence 15, Application US/09509902A
 ; Patent No. 6387676
 ; GENERAL INFORMATION:
 ; APPLICANT: Virco, Duke
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Anderson, Dirk M.
 ; APPLICANT: Marken, John S.
 ; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
 ; FILE REFERENCE: 2877-US
 ; CURRENT APPLICATION NUMBER: US/09/509,902A
 ; CURRENT FILING DATE: 1999-08-03
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 15
 ; LENGTH: 1961
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-509-902A-15

Query Match 24.5%; Score 785.6; DB 4; Length 1961;
 Best Local Similarity 68.6%; Pred. No. 2.1e-204;
 Matches 1098; Conservative 0; Mismatches 499; Indels 3; Gaps 1;

QY 1220 CAGTAAGTAAGCCCTGCAAGATCAACAGCTAGCTTACCCAGAGAGATGATAGTGA 1279
 Db 115 CAGGATGAGCAGAGAGCTCCGTGCGGGCCCAACGAGCCCGAGGGCGGCAAGAGCC 174
 QY 1280 AGAGTGGGACAGTGTGACTCTGGGGATAAAGAAAGAAACAGATTACACTTTCACCTA 1339
 Db 175 GCAAAAGCGCAAGGGGGCAAAAGGAAAGGATAAAAGCAGATTTCAGACCTCTCCCG 234
 QY 1340 TCACTGTCGAAACCAAGAAATCAAGACAGAGAGTCTCGGATTCAGAGTTACT 1399
 Db 235 TTCCCGTCCCAAAACAGCAGACGAGAGAGAGAGTCCAGCATCAAGAGCTACT 294
 QY 1400 CCCACAGCGCTTTAAGTCTTCTATGAGCAGCAGCAGTAGAGACGCTCTCAAGTCTCATC 1459
 Db 295 CGCATACGCTTTAAGTCTTCTATGAGCAGCAGCAGTAGAGATGTCATCAAAACCTACC 354
 QY 1460 AGCAGGCTGTGATCGGAGAGCAGCAGTAGAAAATGAAATGATGCGGGTGGATTTCTC 1519
 Db 355 AGCAGAGGTTAAACCGGAGGCTGAGCTGGAGCAAGAAATGGCAAGCTCTGTG 414
 QY 1520 AAGATGCCAGGATCAAAATGAGAAAGATGCTTTCGCAAGAAAGAGTCTAACTATATTCGTC 1579
 Db 415 AAGTGGAGGAGCAGAGTCCGGAAGATCTCTACCAAGAAAGAGTCTAATTAACACAGGT 474
 QY 1580 TTAAGAGGCTAAATGAGCAAGTCTATGTTTGTAAAGATAAAGACATTAGGAATAGGAG 1639
 Db 475 TAAAGAGGCGCAAGATGGCAAGTCTATGTTTGTCAAGATCAAAACCTCGGGATCGGTG 534

QY	1640	CGTTTGGTGAAGTCGTCTGACGAAGAAAAAGTCGATACATAAGAGTTTGTATGCAACAAAGA	1699
DB	535	CCTTTGGAGAGATGTGCTTTGCTTTGTAAGTGGGACACTCACGCCCTTTTACGCCATGAAGA	594
QY	1700	CTCTTCGAAAGAAAGAGCTTCTGCTCCGAAATCAAGTGGCTCATGTGAAGCGGAGAGGG	1759
DB	595	CCCTAAGGAANAAGGATGTCTCTGAACCGGAATCAGGTGGCCGACGTCAAGGCCGAGAGGG	654
QY	1760	ATATCTCTAGCAGAAGCCGACAATGAGTGGGTGGTCCGCTGTACTACTCTTTTCCAGGACA	1819
DB	655	ACATCTCTGGCGAGGACAGACAATGAGTGGTGGTCAAACTCTACTACTCTCTTCCAAGACA	714
QY	1820	AGGACAACCTTGTACTTTGTGATGGACTACATCTCTGGGGGGGGAATGATGAGCCTATTAA	1879
DB	715	AAGACAGCCTGTACTTTGTGTATGGACTAGATCCCTGGTGGGACATGATGAGCCTGCTGA	774
QY	1880	TTACAATGGGCATCTTTCCCTGANAATCTCGCAGCATTTCTACATACGACAACCTTACCTGTG	1939
DB	775	TCCGGATGGAGGTCTTCCCTTGAGCACTCGCCCGGTTCTACATCGCAGAGTGACTTTGG	834
QY	1940	CAGTTTGAAGTGTTCATAAAATGGGTTTTATTATCAGAGATATTAACCTCATAAACATTT	1999
DB	835	CCATTGAGAGTGTCCACAAGATGGGTTCATCCACCGAGACATCAAGCCTCATAAACATTT	894
QY	2000	TGATTGACCCTGATGGCCATATTAAATTCAGTACTTTTGGCTGTGCATCGCTGGCTTCAGAT	2059
DB	895	TGATAGATCTGGATGGTTCACATTTAAACTCACAGATTTCCGGCTCTGCATCGGTTCAAGT	954
QY	2060	GGACACATGACTCCAAGTACTACCAGAGTGGGGATCCACGAGGCAAGATPAGATGGATT	2119
DB	955	GGACTCACAATCCCAATATTACGAAGAAGGAGGCCATGTCCAGACAGCAGCAGTGGAGC	1014
QY	2120	TCAGTAACGAATGGGAGATCCCTTCCAAATGTTCGTGTGGGACAGACTGAAGCCACTGG	2179
DB	1015	CCACGACCTCTGGATGATGTCTCTAAGTGTTCGTGTGGGACAGCTGAAGACCTTAG	1074
QY	2180	AGCGGAGAGTGTCTGCGCAGCAGCACAGCGATGTCTAGCCCATCTCTGTTGGGACTCCCA	2239
DB	1075	AGCAGAGGGCGGAAGCAGCAGCAGAGGTGCTCGGCAACATTCACCTGTTGACTGTGTGGA	1134
QY	2240	ATTATATTGCACCTGAAGTCTACTCGCAACAGGATATACACAGCTGTGTGACTGTGTGGA	2299
DB	1135	ACTACATCGCACCGAGGTGCTCTCCGCAAAAGGATACACTCAACTCTGTGACTGTGTGGA	1194
QY	2300	GTGTTGTGTATTCTTTGTGAAATTTGTTGGTGGGACAACTCCTTTCTTTGGCACAACCC	2359
DB	1195	GTGTTGGAGTGATCTCTTCGAGATGCTGTGGTGGGCGCGCCCTTTTGGCACCCTACT	1254
QY	2360	CATTAGAAACACAANAATGAAGTTATCATCTGCAAACTTCTCTACATATCCCTCTCTCAAG	2419
DB	1255	CCACAGAAACCCAGCTGAAGGTGATCACTGGGGAACACGCTCCACATTCACGCCCAGG	1314
QY	2420	CTAAGCTGAGTCTCTGAAGCCTCTGACCTCATTTATCAAACCTGTGTGAGGACCAGAAGACC	2479
DB	1315	TGAAGCTGAGCCCTGAGGCCAGGACCTCATACCAAGCTGTGTGCTCGCAGCAGCACCC	1374
QY	2480	GCCTCGCAAGAAGCGTGTGATGAGATAAAGGCTCATCCATTTTAAAGACCATCGATT	2539
DB	1375	GCCTGGGCGGAATGGGGCGATGACCTGAAGGCCACCCCTCTTCTCAGCGCCATGACT	1434
QY	2540	TCCTAGTGATCTGAGACAGCAGTCTGCTTTCATACATCCCTTAAATCACGCATCCCAACAG	2599
DB	1435	TCCTCAGTGACATCCGGAAGCATCCAGCCCCCTAGCTTCCACCATCAGCCACCCCATGG	1494
QY	2600	ATACATCCAATTTTCGACCTGTTGATCTCTGATATAAATTTGTGGAGGATGGCAGCGAGGAGG	2659
DB	1495	ACACCTCGAATTTTCGACCCCGTAGATGAAGAAGGCCCTTTGGAAGGATGCCAGC--GAAG	1551
QY	2660	AAATATACGTGACACTCTCAGCGGATGGTATAAAATGGGAAGCACCACCGGACACGCTT	2719
DB	1552	GTAGCACAAGGCCCTGGGACACACTCACCTCGGCCAATTAACAGCAATCTCTGAGCAGCAT	1611
QY	2720	TCTATGAGTTCACCTTTTCGGAGGTTTTTTTGATGACAATGGCTACCCCATATAATTAATCCAA	2779

D6 1612 TTTACGAATTCACCTTCCGAAAGGTTCTTTTGATGACAATGGCTTACCCTTTTCGATGCCCA 1671

OY 2780 AGCCTATTGACTAGTAATCATTCATTTCACAGGGCTCAGA 2819
||||| | |||||

D6 1672 AGCCTTCAGGAGCAGAGCTTCACAGGCTGAGAGCTCAGA 1711
||||| | |||||

RESULT 5

US-09-509-902A-6

; Sequence 6, Application US/09509902A

; Patent NO. 638766

; GENERAL INFORMATION:

; APPLICANT: Virca, Duke

; APPLICANT: Bird, Timothy A.

; APPLICANT: Anderson, Dirk M.

; APPLICANT: Marken, John S.

; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Fun

; FILE REFERENCE: 2877-US

; CURRENT APPLICATION NUMBER: US/09/509,902A

; CURRENT FILING DATE: 1999-08-03

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 6

; LENGTH: 1498

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-509-902A-6.

Query Match	22.2%	Score 713.8	DB 4	Length 1498	
Best Local Similarity	69.8%	Pred. No. 7.2e-185			
Matches 964	Conservative	0	Mismatches 417	Indels 0	Gaps 1
Qy	1220	CAGTAACTAAGCCCTCGCAAGATGAACACGCTAGCTTACCCAAAGCAAGATGATAGTGAGA	1279		
Db	117	CAGCATGGAGCAGAGCCTCCGTGCGGGCCCCAACGACCCGCGGGCGGCACAGAGACC	176		
Qy	1280	AGAGTGGCGACACTGGTGACTCTGGGGATAAAGAAAACAGCAATTAACAATTCACCTA	1339		
Db	177	GCAAAAGCGCAAGGGGGACAAAGGCGGAAGGATAAAAGACAGATTTCAGACCTCTCCCG	236		
Qy	1340	TCACTGTTCCGSAAAAACAAGAAATGAAGAACGAGAGAGTCTCGGATTTCAGAGTTACT	1399		
Db	237	TTCCGTCGCCAAACACGACAGACGAAGAGAGAGTACGCGATCAAGAGCTACT	296		
Qy	1400	CCCCACAGGCGTTTAAAGTCTTTCATGGAGCAGCAGCTAGAGAACCTCTCGAAGTCTCATC	1459		
Db	297	CGCCATACGCGTTTAAAGTCTTTCATGGAGCAGCAGCTGGAGAAATGTCATCAAAACCTACC	356		
Qy	1460	AGCAGCGTCTGCATCGGAGAGAGCAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTC	1519		
Db	357	AGCAAGAGGTTTAAACCGGAGGCTGACAGCTGGAGCAAGAAATGGCCAAAGCTGGACTGTG	416		
Qy	1520	AAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTGCTC	1579		
Db	417	AGCTGAGCAGGAGCAGATGCGGAAGTCTCTACCAAGAAAGCTCTAATTACAACAGGT	476		
Qy	1580	TAAAGGGCTAAATGGCAAGTCTATGTTTGTAAAGATAAAGACATTAGGAATAGGAG	1639		
Db	477	TAAAGGGCCAAAGATGGACAGTCTATGTTTGTCAAGATCAAAACCCCTGGGGATCGGTG	536		
Qy	1640	CGTTTGGTGAAGTCTCTAGCAAGAAAGTCGATACATAAGCTTTGTATGCAACAAGA	1699		
Db	537	CCTTTGGAGAAGTGTGCCTTGCTTGTGAAGGTGGACACTACAGCCCTGTACGCCCATGAAGA	596		
Qy	1700	CTCTTCGAAAGAAAGACGTTCTGCTCCGAAATTCAGGTGGCTCATGTGAAGCGGAGAGGG	1759		
Db	597	CCCTAAGGAAAAGGATGTCTTGAACCGGAATCAGTGGCCCCAGCTCAAGCCCCGAGAGG	656		
Qy	1760	ATATCTAGCAAGCCGACAAATGAGTGGGTGGTCCGCTCTGTACTACTCTTTTCACGACA	1819		
Db	657	ACATCTTGGCCGAGGACAGACATGAGTGGGTGGTCCAAACTCTACTACTCTCCCAAGACA	716		

QY 1820 AGGACAACTGTACTTGTGATGGACTACATTCCTGGGGGGGATATGATGACCTATTAA 1879
Db 717 AAGACAGCTGTACTTGTGATGGACTACATTCCTGGGGGACATGATGACCTGTGA 776
QY 1880 TTAGAAATGGCATCTTCTCTGAAATCTGGCAGGATTCATACAGCAGAACTTACCTGTG 1939
Db 777 TCCGGATGAGGTCTTCTCTGAGCAGCTGGCCGGTTCATACATCGCAGAGCTGACTTTGG 836
QY 1940 CAGTTGAAGTGTTCATAAATGGGTTTTATTCATAGAGATATTAACTCTGATACATTT 1999
Db 837 CCATTGAGAGTGTCCACAAAGATGGCTTCATCCAGCAGAGATCAAGCTTGATACATTT 896
QY 2000 TGATTGACCGTATGGCATATTAATTTGACTGATCTTTGGCTTGTGCACTGGCTTCAGAT 2059
Db 897 TGATAGATCTGGATGGTTCACATTAATCACTACAGATTTTGGCTCTGCACTGGGTTCAGT 956
QY 2060 GGACACATGACTCCAAAGTACTACAGAGTGGGATCAACCCAGGCAAGATAGCATGGATT 2119
Db 957 GGACTCAAAATTCAAATATTACAGAAAGGAGCCATGTACAGAGCAGCATGGAGC 1016
QY 2120 TCAGTAACGAATGGGAGATCTTCCAAATGTCTGGTGTGGGACAGACTGAAGCCACTGG 2179
Db 1017 CCAGGAGCTCTGGATGATGTGTCTAACTGTGGTGTGGGACAGCTGAAGACCCCTAG 1076
QY 2180 AGCGAGAGTGTCTGCGCAGCAGCAGGATGTCTAGCCCATTTCTTGGTGGGACTCCCA 2239
Db 1077 AGCAGAGGGCGGGAAGCAGCAGAGGTGCTTGGCAGATTCACCTGGTGGGACTCCAA 1136
QY 2240 ATTATATTGACCTGAAGTGTCTGCGAAGAGATATACACAGCTGTGTGAGTGGTGA 2299
Db 1137 ACTACATCGCACCGGAGTGTCTCTCGCAAGAGGTACACTCAACTCTGTGACTGGTGA 1196
QY 2300 GTGTGTGTGTATTTCTTGTGAATGTGTGGGCAAGCTCTTCTTGGGCAAAACC 2359
Db 1197 GTGTGGAGTGTCTTCTGAGATGCTGTGGGGCAGCGCCCTTTTGGCACCCTACTC 1256
QY 2360 CATTAGAACACAAATGAAGTTATCATCTGGCAAACTTCTCTACACATCCCTCTCTCAAG 2419
Db 1257 CCACAGAAACCCAGCTGAAGTGTATCACTGGGAGAACAGCTCCACATTCAGCCCAAG 1316
QY 2420 CTAAGCTGAGTCTGAAGCTCTGACCTCATTAATCAACTGTGTGAGGACAGAACACC 2479
Db 1317 TGAAGTGAAGCTGTAGGCGGAGGAGCTCATCAAGCTGTGTCTCCGACAGCACC 1376
QY 2480 GCCTCGGCAAGACGGTGTGTAGATGAAGGCTCATCCATTTTAAAGACCATCGATT 2539
Db 1377 GCCTGGGCGGAATGGGCGGATGACCTGAAGGCGCCACCCCTTCTTACGCCCATTTGACT 1436
QY 2540 TCTCTAGTGTATGACAGCAGCTGTCTTATACATCCCTTAAATCAGCATCCCAACAG 2599
Db 1437 TCTCCAGTGCATCCGGAAGCATCCAGCCCTTAGGTTCCCAACCATCAGCACCCTATGG 1496
QY 2600 A 2600
Db 1497 A 1497

RESULT 6
US-09-442-100-1
; Sequence 1, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weli
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; NUMBER OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas

CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistock, S. Leslie
REGISTRATION NUMBER: 16,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 5720 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA
NAME/KEY: CDS
LOCATION: 1103..4402
US-09-442-100-1

Query Match 16.2%; Score 521.2; DB 4; Length 5720;
Best Local Similarity 65.4%; Pred. No. 5.5e-132;
Matches 852; Conservative 0; Mismatches 423; Indels 27; Gaps 5;

QY 1354 AACAGAAAGATCAAGAACAGAGAGTCTCGGATTCAGAGTTACTCCACAGAGCCTTT 1413
Db 2984 AAGGAGAGGAGGAGGAGGAGGAGGAGTCTCCGATCAGGAGTACTCGCGCAAGCCTTC 3043
QY 1414 AAGTCTTCTATGAGCAGCAGCTAGAGAACGTCCTGAAGTCTCATCAGCAGGCTGTGAT 1473
Db 3044 AAGTCTTCTATGAGCAGCAGCAGTATAGAACGTCGTATCGGATTCGCGCAGCAGTAT 3103
QY 1474 CGCAAGAACAGCTAGAAATGAAATGATGCGGGTGGATTATCTCAAGATGCCAGGAT 1533
Db 3104 CGCAAGAACAGCTAGAACGAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCAAATC 3163
QY 1534 CAAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCGTCTTAAAGGGCTAAA 1593
Db 3164 GAGATCAGAAATGCTGAACCAAGGAGAGCAACTACATTCGATTGAAGCGGCCAAG 3223
QY 1594 ATGGCAAGTCTATGTTTCTAAAGATAAAGACATTAGGAATAGGAGCGTTTGGTGAAGTC 1653
Db 3224 ATGGCAAGAGCATGTTCTGCTCAAACTGAAGCCCATTTGAGTGGTGTGATTTGGCAGGTA 3283
QY 1654 TGCTCTAGCAAGAAAGTCGATA---CTAAAGCTTTGTATGCAACAAAGACTCTTCGAAAG 1710
Db 3284 AGCTGTGGTGAACAAATCGATACCTCGAACCAATTTGTATGCGATGAACCCCTCGCGAAA 3343
QY 1711 AAAGAGTCTCTCGCAAAATCAGGTGGCTCATGTGAAGCGGAGAGGATATCTCTAGCA 1770
Db 3344 CGGACGCTCTCAAGCGCAATCAGGTGGCAGACGTCGAGCCGAGAGGATATCTCTCGCG 3403
QY 1771 GAAGCCGACATAGTGGTGGTCCGCTGTACTACTCTTCTTCCAGGACAAAGCAACTTG 1830
Db 3404 GAAGCCGACAAATAGTGGTGGTGAAGTTTGTACTACTACAGCTTCCAGGACAAAGGATAATCTG 3463
QY 1831 TACTTTGTGAGTACTACATTCCTGGGGGGGATATGATGACCTATTATTAAGATGGGC 1890

```

Db 3464 TACTTTGATGGACTACATACAGGTGGTGTGATCGCTGCTCATCAAACTGGGC 3523
QY 1891 ATCTTCTCTGAAATCTGCGCAGATCTACATAGCAGAACTTACCTGTGCATTTGAAAGT 1950
Db 3524 ATTTTGGAGGAGAACTGCGCAGATCTACATCGCGGAGGTACCTGGCCCGTGGACAGC 3583
QY 1951 GTTCAATAAATGGTTTATTCATAGAGATATTAACCTGATAAGATTTTGTGATGACCGT 2010
Db 3584 GTTCAATAAATGGTTTATTCATAGAGATATTAACCTGATAAGATTTTGTGATGACCGT 2010
QY 2011 GATGCCATATTAATTTGACTGACTTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2070
Db 3644 GACGACACATAAAGCTACCGACTTTGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3703
QY 2071 TCCAACTACTACC---AGAGTGGGATCACCCACGCGAAGATAGATGATGATTTAGTAC 2127
Db 3704 TCGAAGTACTACAGAGAACGCGAATCACTCGCCGAGGACTCGATGAGGCCCTGGGAG 3763
QY 2128 GAATGGGAGATCTTCCAAATTTGCGGTGTGGGACAGACTGAAGCCACTGGAGCGGAGA 2187
Db 3764 GAAT-----ACTCCGAGAAGGACGAGCCACCCACCGTGTGGAGAGCGA 3808
QY 2188 GCTGCTCGCCACACAGGATGCTAGCCCAATCTCTGTTGGGACTCCCAATATATTT 2247
Db 3809 CGGATGCGCGATACCAAGAGTCTGCGCCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3868
QY 2248 GCACCTGAAGTCTACTGCGAAGAGATATACACAGCTGTGTGACTGTGGTGGAGTCTTGGT 2307
Db 3869 GCTCCGAGGCTCTGGAGAGAGTGGGTACACGCGAGCTGTGCGACTACTGGAGCGTGGGC 3928
QY 2308 GTTATCTTTGAAATGTTGGTGACAACTCTCTTTGGCACAACCCCACTTAGAA 2367
Db 3929 GTCATCTTTACGAGATGCTGTGGTGTGCGCGCCCTTTCTGGCCACACAGTCCGCTGGAA 3988
QY 2368 ACACAATGAAGTTATCATCTGGCAAACTCTTACATCCCTTCTACATCCCTTCTCAAGCTAAGCTG 2427
Db 3989 AGCAACAAGAGTCTCAACTGGGAGAAACGCTGCAATATTCGCGCCGACGCGGAGTTA 4048
QY 2428 AGTCTGAAGCCTCTGACTCATATTAATCAACTGTGTGAGGACCAAGACCGCTCGGC 2487
Db 4049 TCCCGGAGGCTACGAGTGTATAGGAGGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4108
QY 2488 AGAAGCTGTGTATGAGATAAGGCTCATCCATTTTAAAGACATCGATTTCTCTAGT 2547
Db 4109 AAGA---GCGTGGACGAGTCAAGAGCCAGCACTTCTCAAGGGCATCGA---CTTTGGC 4162
QY 2548 GATCTGACACAGCTGTCTTCATACATCCCTTAAATCAAGCATCCACAGATACATCC 2607
Db 4163 GACATGCGGAAGCAAGAAAGCGCCCTACATACCGGAAATCAAGCACCCCAAGCAGACATCC 4222
QY 2608 AATTTGACCCCTGTTGATCGATGATAAATTTGTGAGCGGATGGC 2649
Db 4223 AACTTTGATCCGCTGGATCCCGGAGAAGCTGCGCTCGAATGAC 4264

```

RESULT 7

US-09-328-111-26/C

; Sequence 26, Application US/09328111

; Patent No. 6262333

; GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.

; APPLICANT: Steinmann, Kathleen E.

; APPLICANT: Astle, Jon H.

; APPLICANT: Burgess, Christopher C.

; APPLICANT: Bushnell, Steven E.

; APPLICANT: Carroll III, Eddie

; APPLICANT: Catino, Theodore J.

; APPLICANT: Derti, Adnan

; APPLICANT: Ford, Donna M.

; APPLICANT: Lewis, Marcia E.

; APPLICANT: Monahan, John E.

; APPLICANT: Schlegel, Robert

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; FILE REFERENCE: CDD-257 (US)

; CURRENT APPLICATION NUMBER: US/09/328,111

; EARLIER FILING DATE: 1999-06-08

; EARLIER APPLICATION NUMBER: US 60/088,801

; EARLIER FILING DATE: 1998-06-10

; NUMBER OF SEQ ID NOS: 850

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 26

; LENGTH: 638

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-328-111-26

Query Match 16.1%; Score 518.6; DB 4; Length 638;

Best Local Similarity 89.5%; Pred. No. 8.3e-132;

Matches 570; Conservative 0; Mismatches 64; Indels 3; Gaps 1;

```

QY 1444 GTCCTGAAGTCTCATCAGCAGCGTCTGCATCGGAAGAAGCAGCTAGAAAATGAAATGATG 1503
Db 637 GTACTCAATCTCATCAGCAGCGTCTACATCGTAAAAACAATTAGAAGTAATGATG 578
QY 1504 CGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATCTTTGGCCAGAAAG 1563
Db 577 CGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATCTTTGGCCAGAAAG 518
QY 1564 TCTACTATATCTCTTAAAGGCTTAAATGGACAGTCTATGTTTGAAGATAAG 1623
Db 517 TCTAATTCATCCGCTTAAAGGCTTAAATGGACAGTCTATGTTTGAAGATAAG 458
QY 1624 ACATTAGGAATAGGAGCGTTGGTGAAGTCTGTCTAGCAAGAAAGTGCATCTAAAGCT 1683
Db 457 ACATTAGGAATAGGAGCGTTGGTGAAGTCTGTCTAGCAAGAAAGTGCATCTAAAGCT 398
QY 1684 TTGTTATGCAAAAGACTCTTTCGAAAGAAAGACGTTTCTGCTCCGAAATCAGGTGCTCAT 1743
Db 397 TTGTTATGCAAAAGACTCTTTCGAAATAAAGATGTTCTTCTCGAAATCAAGTCTCAT 338
QY 1744 GTGAAGCGGAGAGGATATCTTAGCAGACCGCAATAGTGGTGGTGGTGGTGGTGGTGGT 1803
Db 337 GTTAAGCGTGAAGAGATATCTTGGCTGAAGTGAAGTGAAGTGAAGTGAAGTGAAGTGAAG 278
QY 1804 TACTCTTTCCAGGCAAGGACAACTTGTACTTTGTGATGGACTACATCTCTCTGGGGGGAT 1863
Db 277 TATTCAATCCAGATAAGGACAAATTTATCTTTGTAATGGACTACATCTCTCTGGGGGGAT 218
QY 1864 ATGATGCGCTTATTAATTAAGTGGCATCTTCTGCTGAAATCTGGCAGATTTCTACATA 1923
Db 217 ATGATGCGCTTATTAATTAAGTGGCATCTTCTGCTGAAATCTGGCAGATTTCTACATA 158
QY 1924 GCAGAACTTACCTGTCAGTTCAAAAGTG---TTCAATAAATGGGTTTATTCATAGAGAT 1980
Db 157 GCAGAACTTACCTGTCAGTTTGAAGGTTGCTCTTAAATGGGTTTATTCATAGAGAT 98
QY 1981 ATTAACCTGATAACATTTTGAATGACCGTGTATGATGATGATGATGATGATGATGATGAT 2040
Db 97 ATTAACCTGATAATATTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 38
QY 2041 TTGTGACTGGCTTCAGATGGACACATGACTCCAAAGT 2077
Db 37 CTCTGACTGGCTTCAGATGGACACACGATTTCTAAGT 1

```

RESULT 8

US-09-328-111-66

; Sequence 66, Application US/09328111

; Patent No. 6262333

; GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.

; APPLICANT: Steinmann, Kathleen E.

; APPLICANT: Astle, Jon H.

; APPLICANT: Burgess, Christopher C.

	Matches	500;	Conservative	0;	Mismatches	97;	Indels	10;	Gaps	6;
QY	1447	CTGAAGTCTCATCAGCAGCGTCTGCATCGGAAGAGCAGCTAGAAAAATGAATGATCGG	1506							
Db	2	CTCAAAATCTCATCAGCAGCGTCTACATCGTAAAAACAATTAGAGAATGAATGATCGG	61							
QY	1507	GTTGGATTATCTCRAGATGCCAGATCAATGAGAAAGATGCTTTGCCAGAAAGAGTCT	1566							
Db	62	GTTGGATTATCTCRAGATGCCAGATCAATGAGAAAGATGCTTTGCCAAAAAGAAATCT	121							
QY	1567	AACATATATCTCTTAAAGGGCTTAAATGACAAGTCTATGTTTGAAGATAAAGACA	1626							
Db	122	AATTACATCCGCTTAAAGGGCTTAAATGACAAGTCTATGTTTGAAGATAAAGACA	181							
QY	1627	TTAGGAATAGAGCGTTTGGTGAAGTCTGTCTAGCAGAAAAGTCGATCTAAAGCTTTG	1686							
Db	182	CTAGGAATAGAGCATTTGGTGAAGTCTGTCTAGCAGAAAAGTAGATCTAAGGCTTTG	241							
QY	1687	TATGCAACAAGACTCTTCGAAAGAAAGAGCTTCTGCTCCGAAATCAGTGGGTCTCATGTG	1746							
Db	242	TATGCAACAACAACCTCTTCGAAAGAAAGAGTGTCTTCTTCGAAATCAAGTCGCTCATGTT	301							
QY	1747	AAAGCGGAGAGGATATCCCTAGCAGAAAGCGCACAAATGAGTGGGTGGTCCGCTGTACTAC	1806							
Db	302	AAGGCTGAGAGAGATATCCCTGGCTGAAGCTGACAAATGAATGGGTAGTTCGCTCATATAT	361							
QY	1807	TCCTTCCAGGACAGGACAACTTGTACTTTGTGATGGACTACATTCCTGGGGGGATATG	1866							
Db	362	TCATTCAGATAGGGCCATTATCCCTT--GTAATGGCTACATTCCTNGGGGTGATATG	419							
QY	1867	ATGAGCCTATTAATAGAATGGCATC-TTTCCTGAAAACTCGGCACGATTCCTACAT--A	1923							
Db	420	AGAGGCCATTAATTAATAGATGGCATCTTTCCAGAAAGCTNGCACCATCTACCTTAG	479							
QY	1924	GCAGAACTTTACCTGTG-CAGTTGAAGTCTTCAT--AAAATGGGTTTTATTCATAGAGAT	1980							
Db	480	CCAGAACTTTACCTGNCNGTTGAAAGTGGTCTTTAAATAGGGGTTAATCTTACAGAT	539							
QY	1981	ATTAACCTTGATAACATTTTGATG--ACCGTATGGCCATTAATAATGACTGACTTTG	2038							
Db	540	TTTATACCTGGATAATATTTGANTGGACGNGAAGGCCCTTATTAATAATGGCTTGCTTG	599							
QY	2039	CGTTGTG	2045							

```

Db 302 CTGCTTGGATGACTTTGAGTCTCTGAAGATTATAGGAAGAGGAGCTTTTGGAGAGTG 361
Qy 1654 TGTCTAGCAAGAAAGTCTGATCTAAAGCTTTGTATGCAACAAAGACTCTTCTGGAAGAAA 1713
Db 362 CGTGTGGTCCACAAAAGATACAGGCCATATCTATGCAATGAAGATATTGGAAGTCT 421
Qy 1714 GACGTTCTGCTCCGAAATCAGTGGCTCATGTGAAGCGGAGGAGATATCTTACAGAA 1773
Db 422 GATATGCTTGAAGAAAGACAGTGGCCCATATCCAGCAGAAAGAGATATTGTTGTAGAA 481
Qy 1774 GCCGACATGAGTGGTGGTCCGCTGTACTACTCTTTCCAGGACAAGCAACTGTGAC 1833
Db 482 GCAGATGCTGGTGGTGGTGAAGATGTTTACAGTCTTTCAGGATTAAGAGGAATCTTAT 541
Qy 1834 TTTGTGATGACTACATCTCTGGGGGATATGATGACCTTATTAATTAGATGGGCATC 1893
Db 542 CTATCATGGAATTTCTCCCTGGAGTGACATGATGACATCTGCTAATGAAGAAAGACACC 601
Qy 1894 TTTCTGAAATCTGGCAGCATTTACATAGCAGAACTTACCTGTGCAAGTTGAAAGTGT 1953
Db 602 TTGACAGAAGAGGAACACAGTCTTCTACATTTTCAGAGACTGTTCTGCAATAGATCGCATC 661
Qy 1954 CATAAATGGTGTATTCATAGATATTAACCTGTATACATTTTGATGACCGTGAT 2013
Db 662 CACCAGTTGGTGTTCATCCATCGGATATTAAGCCAGACAACTTTTATGGATGCCAAG 721
Qy 2014 GGCATATTAATGACGTACTTTGGCTTTGTCAGTGGCTTCAGATGACACATGACATCC 2073
Db 722 GGTATGTAATATCTGATTTTGGTTCATGTCAGGATTAAGAAAGCTCACAGGACT 781
Qy 2074 AAGTACTACAGAGTGGGATCACCACGCAAGATAGCATGATTTTCAGTAACGAATGG 2133
Db 782 GAATTTTAT-----AGAAATCTCACACACACACCAACCAAGTG 818
Qy 2134 GGAGATCTTCAATTTGCTGGTGGGACAGATGAAGCCACTGAGCGGAGAGCTGCT 2193
Db 819 ACTTCTCATTTTCAGAACATGAATCAAGAGAGAAAGCAAACTTTGGAAGAAGA----- 872
Qy 2194 CGCCAGCACACGAGTCTAGCCCATCTCTGGTGGGACTCCCAATATATGTCACCT 2253
Db 873 -----ACAGAGCAACTGGCATATTTCCAGTGTGGACACAGATATGCTCTCA 925
Qy 2254 GAAGTGTCTACGCAACAGGATATACACAGCTGTGCTGAGTGGTGGTGGTGTAT 2313
Db 926 GAAGTATTCATGACAGCTGTTTACAACTTTGCTGACTGCTGCTTTGGGACTGAT 985
Qy 2314 CTTTGTGAATGTTGGTGGGACACCTCTCTTTTGGCACAACCCCATTAAGAACACAA 2373
Db 986 ATGATGAATGCTAATAGGATATCCACCTTTCTGCTCTGAAACACCTCAAGAAACATAC 1045
Qy 2374 ATGAAGTTATCATCTGCAAACTTCTTACACATCTCCCTCTCAAGCTAAGCTGATCCT 2433
Db 1046 AGAAAGTGTGACTGGAAGAACTCTGATTTCTCCAGAGGTACCTATATCTGAG 1105
Qy 2434 GAAGCTCTGACCTCATTTATCAAACTGTGTGAGAGACAGAGCGCTCGGCAAGAAC 2493
Db 1106 AAAGCAAGGACTTAATCTCAGATTTTGTATGATTTCTGAAACAGAAATTTGGAATAGT 1165
Qy 2494 GGTCTGATGATGAAGAGCTCATCCATTTTAAAGACCCTGAT 2539
Db 1166 GGAGTAGAAGAAATAAAGGTCTATCCCTTTTGAAGGTGTCGACT 1211

```

RESULT 10

US-09-272-796-11
 ; Sequence 11, Application US/0927296
 ; Patent No. 6207148
 ; GENERAL INFORMATION:
 ; APPLICANT: Bandman, Olga
 ; APPLICANT: Hillman, Jennifer L.
 ; APPLICANT: Corley, Neil C.
 ; APPLICANT: Guegler, Karl G.

```

; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; TITLE OF INVENTION: KINASES
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J.
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAT04
; CLONE: 705365
; US-09-272-796-11

```

```

Query Match 7.7%; Score 248.4; DB 4; Length 1935;
Best Local Similarity 53.2%; Pred. No. 8.3e-58;
Matches 599; Conservative 0; Mismatches 491; Indels 36; Gaps 2;

Qy 1414 AAGTTCTTTCATGGAGCAGCAGTAGAGAACGTCCTCTGAAAGTCTCTATCATCAGCAGCGTCTGCAT 1473
Db 122 AAGCTCACATTGGAGAAATTTTATAGCAACCTAATTTTACAGCATGAAGAGAGAAACC 181
Qy 1474 CGAAGAAGCAGCTAGAAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGGAT 1533
Db 182 AGGCAGAGAAATTTAGAAAGTGGCCATGGAAGAAGAGATTAGCAGATGAAGAGAAAAAG 241
Qy 1534 CAATGAGAAAGATGCTTTGGCCAGAAAGCTTAATACTATTTTCTGCTTTTAAAGGGCTAAA 1593
Db 242 TTACGTCGATCACACACGCTCGCAAGAAACAGAGTTCTTACGGCTCAAAAGGACACAGA 301
Qy 1594 ATGGAAGAGTCTATGTTTGTAAAGATAAAGACATTAGGAATAGGAGCGTTTGGTGAAGTC 1653
Db 302 CTGCGTTGGATGACTTTTGGAGTCTCTGAAAGTTATAGGAAGAGAGGCTTTTGGAGAGGTG 361
Qy 1654 TGTCTAGCAAGAAAGTGCATATCACTAAAGCTTTTGTATGCAACAAAGACTCTTCTGAAAGAA 1713
Db 362 CGGTTGGTCCCAAAAAAGATACAGGCCATATCTATGCAATGAAGATATTGAGAAGTCT 421
Qy 1714 GACGTTCTGCTCCGAAATCAGTGGCTCATGTGAAGCGGAGAGGATATCTTACAGAA 1773
Db 422 GATATGTTGAAAGAGACAGGTTGGCCCATATCCAGCAGCAAGAAAGAGATTTTGTGTAGAA 481
Qy 1774 GCCGACATGAGTGGTGGTGGTCCGCTGTACTACTCTTCCAGGACAAGCAACTGTGAC 1833

```

FEATURE:		NAME/KEY: CDS		Score 244.4; DB 2; Length 3018;	
US-08-860-150-6		LOCATION: (596)...		Pred. No. 13e-56;	
		Query Match		Mismatches 491; Indels 39; Gaps 3;	
		Best Local Similarity		Matches 608; Conservative	
QY	1478	AGAAGCAGCTAGAAATCAATGATCGGGTGTGATATATCTCAAGATGCCAGGATCAAA	1537		
DB	729	AAAAGAGTTAGAAAGGTGATGAAGAAGAGCCCTAAAGATGAGAGAGAACGACTCC	788		
QY	1538	TCAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCGCTTTAAAGGGCTAAAATGG	1597		
DB	789	GGGATCAGCAGCATGCTCGGAAGAAACAGAGTTCTTCGTTTGAAGAGAAACAGACTTG	848		
QY	1598	ACAAGTCTATGTTTGAAGATTAAGACATTAAGATAGGAGCGTTTGGTGAAGTCTGTCT	1657		
DB	849	GATTGGAAGATTTTGGTCTTAAAGTAAATAGGCAGAGGAGCATTTTGGTGAAGTACGCC	908		
QY	1658	PAGCAAGAAAGTCCGATACATAAGCTTTGTATGCAACAAAGACTCTTCAAGAAAGACG	1717		
DB	909	TTGTTCAAGAAAGATACGGGACATGTATGCAATGAAATACTCCGTAAGACAGATA	968		
QY	1718	TTCTGCTCCGAAATCAGGTGCTCATGTGAAGCGGAGGAGGATATCTTAGCAGAACCG	1777		
DB	969	TGCTTGAAGAAAGAGAGGTGGCCACATTCGTGCGGAGCGTGACATTTAGTGGAGGCG	1028		
QY	1778	ACAATGAGTGGGTGGTCCGCTGTACTCTTCCAGGACAAGGACAACACTTGTACTTTG	1837		
DB	1029	ACAGTTTGGGTGTGAAATGTTCTATGTTTTCAGGATAAGCTAAACCTCTACCTAA	1088		
QY	1838	TGATGACTACATTCCTGGGGGGATATGATGAGCTATTAATAGATGGGATCTTTC	1897		
DB	1089	TCATGAGTTCCTGCTGCGGGGACATGATGACCTTGTGATGAAAAGACACTCTGA	1148		
QY	1898	CTGAAATCTGGACCATCTACATAGCAGAACTTACCTGTGACGTTGAAAGTGTTCATA	1957		
DB	1149	CAGAAGAGGAGACTCATGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATTCC	1208		
QY	1958	AAATGGTGTATATATAGATATTAACCTGATTAACATTTTGTATGACCGTGATGGCC	2017		
DB	1209	AACTTGGATTCATCCACAGAGACATCAACACAGACAACCTCTCTTTTGGACAGAGGCC	1268		
QY	2018	ATATTAATGACTGACTTTGGCTTGTGCACTGGCTTCAGTGGACACATGACTCCAACT	2077		
DB	1269	ATGTGAACTTCTGACTTTTGGTCTTTGACAGGACTGAAAAGACACATAGCAGAAAT	1328		
QY	2078	ACTACAGAGTGGGATCACCCAGCAAGATAGCATGATTTTCAGTAACGAATGGGAG	2137		
DB	1329	TTTATAGAACTCTGAAACACAGCTCCCACTGATTTTCACTTTCCAGAAC	1378		
QY	2138	ATCCTTCCAATTTGCTGGTGGGACAGACTGAAGCCACTGGAGCGGAGAGCTGCTGCC	2197		
DB	1379	-----ATGAATTCCTCAAGAAAGAGAGAACTCTGGAAGAG	1413		
QY	2198	AGCAGCAGGATGCTAGCCCATCTCTGTGGGACTCCCAATTTATTTGCACTGAAG	2257		
DB	1414	A-AATAGACGCTGACAGCTCTTCCACAGTAGGCACTCTCTGACTACATTTGCTGAGG	1472		
QY	2258	TGCTACTCGCAACAGGATATACACAGCTGTGTGACTGGTGGAGTGTGTTTATCTTT	2317		
DB	1473	TGTTCAATGCAGACCGGTGACAAAGCTCTGTGATTTGGTGTGCTTTGGGTGATCATGT	1532		
QY	2318	GTGAAATTTGGTGGGCAACCTCTTCTTGGCACAACCCCATTTAGAAACAAATGA	2377		
DB	1533	ATGAGATGCTCATCGGCTACCCACTTCTCTGTGAGACCCCTCAGAGACATATAGA	1592		
QY	2378	AGGTTATCATCTGCAAACTTCTCTACACATCTCTCTCAAGTCAAGTCTCTCTGAAG	2437		
DB	1593	AGGTGATGAACTGGAAAGAACTTTGACTTTTCTCTCCAGAAAGTCTCTCTGAGAAAG	1652		

DB	482	GCAGATGGTGCCTGGTGTGAAGATGTTTACAGTTTTTCAGGATAAGAGGAATCTTAT	541		
QY	1834	TTTGTGATGAGTACATCTCTGGGGGGATATGATGAGCTTATTAATTAAGATGGGATC	1893		
DB	542	CTAATCATGAATTTCTCCCTGGAGTGACATGATGACATGCTATGTAAGAAACACACC	601		
QY	1894	TTTCTCGAAATCTGCAGCATCTACATGAGAGAACTTACCTGTGCTGAGTTGAAAGTGT	1953		
DB	602	TTGACAGAGAGAAACACAGTCTACATTTTACAGACTGTTCTGCAATAGATGCGATC	661		
QY	1954	CATAAAATGGTGTATTCATAGAGATATTAACCTGATTAACATTTTGTGACCGTGTAT	2013		
DB	662	CACAGTTGGTGTTCATCCATCGGATATTAAGCCAGACAACCTTTTATTTGGATGCCAAG	721		
QY	2014	GGCCATTAATTAATGACTGACTTTGGTGTGCTGCACTGGCTTCAGATGGACATGACTCC	2073		
DB	722	GGTCATGTAATTAATCTGATTTTGGTTCATGTACGGGATTAAGAAAGCTCAGAGACT	781		
QY	2074	AAGTACTACCAGAGTGGGATCACCCAGCAAGATAGCATGATTTTCACTAAAGATGG	2133		
DB	782	GAATTTTAT-----AGAAATCTCACACACAACCCCAAGTG	818		
QY	2134	GGAGATCTTCCATTTGCGGTGGGACAGACTGAAGCCACTGGAGCGGAGAGCTGCT	2193		
DB	819	ACTTCTCATTTCAAGACATGAACCTCAAGAGAGAAAGCAGAACTTTGGAAGAAG	872		
QY	2194	CGCCAGCAGCAGGATGCTAGGCCATTTCTGTTGGGACTCCCAATTAATTTGCACT	2253		
DB	873	-----ACAGGACAACTGTCATATTCACAGCTTTGGACACACAGATTTACATTTGCTCA	925		
QY	2254	GAACTGTACTCGCAACAGGATATACAGCTGTGCTGACTGGTGGTGTGTTTAT	2313		
DB	926	GAAATTAATCATCGACACTGGTTACAACAAATTTGCTGACTGGTGGTCTTTGGGAGTATT	985		
QY	2314	CTTTGTGAATTTGGTGGGCAACCTCTTCTTGGCACAACCCCATTAAGAAACAA	2373		
DB	986	ATGATGAATGCTAATAGATATCCACCTTCTGCTGAAACACCTCAAGAAACATAC	1045		
QY	2374	ATGAAGTTATCATCTGCAAACTTCTTACACATCCCTCTCAAGCTAGTCCCT	2433		
DB	1046	AGAAAGTGAATGAACTGGAAGAACTCTGATTTCTCAGAGGTAACCTATATCTGAG	1105		
QY	2434	GAAGCTCTGACCTCATATCAAACTGTCTGAGGACCAAGACCGCTCGGCAAGAAC	2493		
DB	1106	AAAGCCAGGACTTAATCTCAGATTTTGTATTGATTTGAAACAGAAATTTGGAATAGT	1165		
QY	2494	GGTGTGATGAGATAAGGCTCATCCATTTTAAAGACCATCGATT	2539		
DB	1166	GGAGTAGAAGAAATAAAGTCTATCCCTTTTGAAGGTGTCGACT	1211		

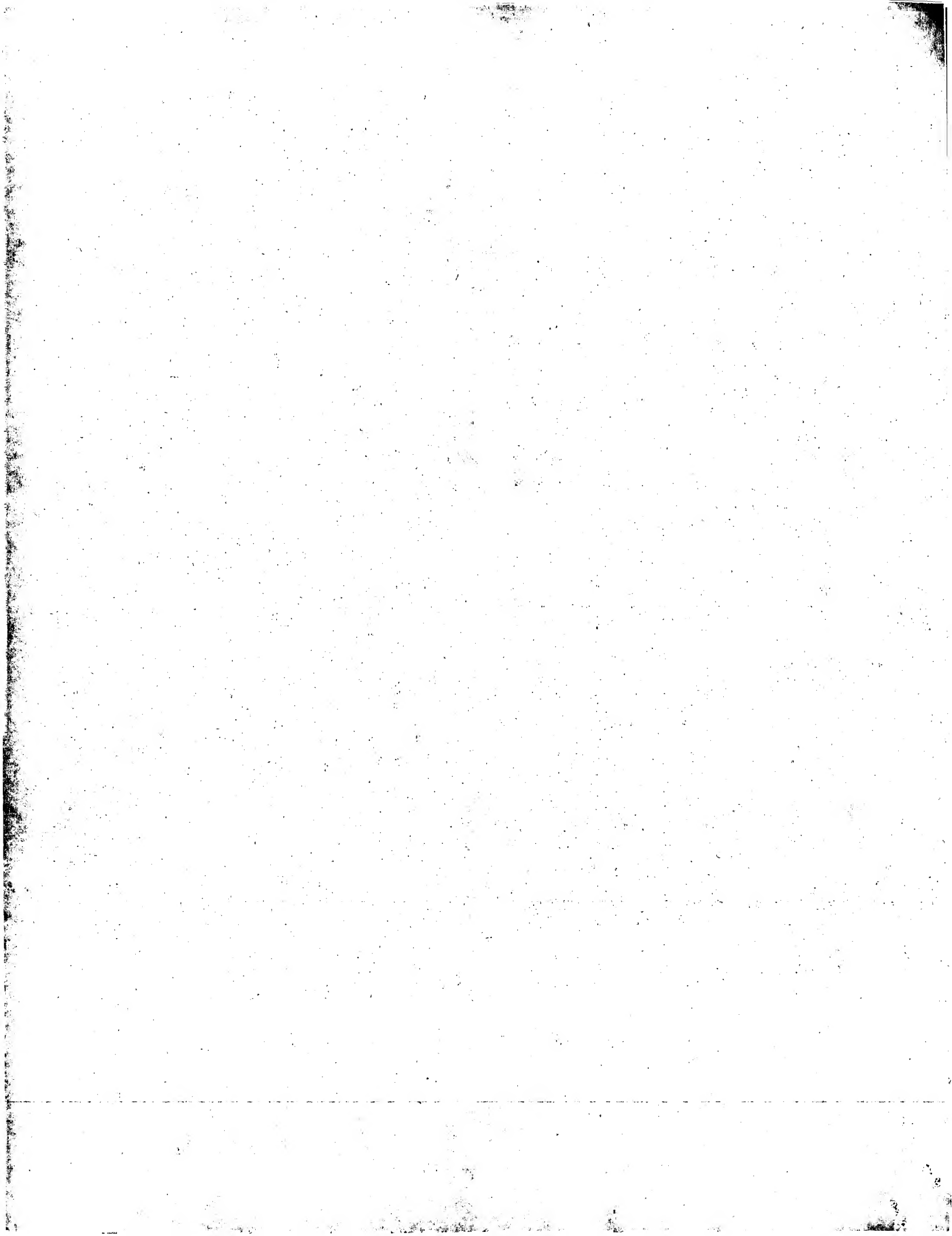
RESULT 11
US-08-860-150-6
Sequence 6, Application US/08860150B
Patent No. 5981205
GENERAL INFORMATION:
APPLICANT: Hemmings, Brian A.
APPLICANT: Millward, Thomas A.
TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
FILE REFERENCE: 4-20265/A/PCT
CURRENT APPLICATION NUMBER: US/08/860,150B
CURRENT FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: PCT/EP95/05052
EARLIER FILING DATE: 1995-12-20
EARLIER APPLICATION NUMBER: 94810746.1
EARLIER FILING DATE: 1994-12-22
NUMBER OF SEQ ID NOS: 18
SOFTWARE: PatentIn Ver. 2.0
SEQ ID NO 6
LENGTH: 3018
TYPE: DNA
ORGANISM: Homo sapiens

[illegible]

RESULT 15
US-09-599-256-1.

	Query Match	5.6%	Score 179.8;	DB 4;	Length 2160;
	Best Local Similarity	49.8%;	Pred. No. 4.6e-39;		
	Matches 530;	Conservative	0;	Mismatches 487;	Indels 48; Gaps 1;
QY	1509	TG GATTATCTCAAGATGCCCGAGATCAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAA	1568		
DB	828	TGCGTGCTCTGAAGAGAGAAGAAGACACAGACAGCTTTCTTCACTTGGGTAAAAAGGATCGCA	887		

Search completed: January 16, 2003, 21:56:16
Job time : 124.372 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:42 : Search time 89.1612:Seconds
(without alignments)
16067.254 Million cell updates/sec

Title: US-09-763-334-3
Perfect score: 3213
Sequence: 1 gtgcacattcaattacgc.....cccccaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 393868 seqs, 22934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:**
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Length	DB ID	Description
1	681	21.2	2043	Sequence 2, Appli
2	518.6	16.1	638	Sequence 26, Appli
3	376.6	11.7	678	Sequence 66, Appli
4	371.8	11.6	676	Sequence 214, App
5	247.6	7.7	1689	Sequence 1861, Ap
6	244.4	7.6	3583	Sequence 152, App
7	212.6	6.6	1452	Sequence 2402, Ap
8	184	5.7	1818	Sequence 89, Appli
9	175.2	5.5	734	Sequence 196, App
10	136.4	4.2	1416	Sequence 2503, Ap
11	132.8	4.1	1244	Sequence 38, Appli
12	132.8	4.1	1393	Sequence 37, Appli
13	132	4.1	1735	Sequence 58, Appli
14	130.2	4.1	1398	Sequence 633, App
15	117.4	3.7	2365	Sequence 6, Appli
16	115.8	3.6	2311	Sequence 6, Appli
17	115.8	3.6	2370	Sequence 1, Appli
18	115.8	3.6	2370	Sequence 214, App
19	115.8	3.6	2370	Sequence 3855, Ap

20	113	3.5	1515	10	US-09-804-471A-1	Sequence 1, Appli
21	113	3.5	5877	12	US-10-028-946-3	Sequence 3, Appli
22	113	3.5	6165	12	US-10-028-946-1	Sequence 1, Appli
23	112.8	3.5	1479	10	US-09-771-161A-46	Sequence 46, Appli
24	111.4	3.5	2637	10	US-09-799-875-3	Sequence 3, Appli
25	111.4	3.5	3003	10	US-09-799-875-1	Sequence 1, Appli
26	109.8	3.4	6159	9	US-10-017-216-3	Sequence 3, Appli
27	109.8	3.4	6574	9	US-10-017-216-1	Sequence 1, Appli
28	109.2	3.4	277	10	US-09-294-093B-2043	Sequence 2043, Ap
29	109	3.4	568	9	US-09-796-692-8122	Sequence 8122, Ap
30	109	3.4	5313	10	US-09-801-368-297	Sequence 297, App
31	109	3.4	6409	10	US-09-864-864-293	Sequence 293, App
32	107	3.3	2549	10	US-09-880-107-3691	Sequence 3691, Ap
33	105	3.3	3407	10	US-09-971-845-1	Sequence 1, Appli
34	103.8	3.2	587	9	US-09-764-868-217	Sequence 217, App
35	100	3.1	1194	10	US-09-801-368-399	Sequence 399, App
36	97.2	3.0	2706	12	US-10-071-751-20	Sequence 20, Appli
37	96.4	3.0	2261	9	US-09-954-531-141	Sequence 141, App
38	96.4	3.0	2261	9	US-09-954-531-570	Sequence 570, App
39	96.4	3.0	3061	10	US-09-880-107-2146	Sequence 2146, Ap
40	96.2	3.0	362	9	US-09-796-692-7789	Sequence 7789, Ap
41	96	3.0	6303	9	US-09-974-298-168	Sequence 168, App
42	95.4	3.0	1635	10	US-09-880-107-2340	Sequence 2340, Ap
43	94.2	2.9	512	10	US-09-867-701-6230	Sequence 6230, Ap
44	90	2.8	1197	10	US-09-801-368-403	Sequence 403, App
45	89.8	2.8	2698	10	US-09-749-956-3	Sequence 3, Appli

ALIGNMENTS

RESULT 1
US-09-836-392-2
; Sequence 2, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 2
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-2

Query Match	21.2%	Score	681;	DB	9;	Length	2043;
Best Local Similarity	70.4%	Pred. No.	4.4e-159;				
Matches	925;	Conservative	0;	Mismatches	385;	Indels	3;
Gaps	1;						
QY	1507	GTTCGATTATCTCAAGATGCCCGGCTAAATGAGAAAGATGCTTTGCCAGAAAGAGTCT	1566				
Db	218	GCTGCACTCTGTGAAGCTGAGCAGGAGGAGATGCGGAAGATCCTCTACGAAAGAGTCT	277				
QY	1567	AACTATATTCGCTTAAAGGGCTAAATGGAACAAGTCTATCTTTGTAAGATAAGACA	1626				
Db	278	AATTACACAGCTTAAAGAGGCCGAGATGAGCAAGTCTATCTTTGTCAGATCAAAACC	337				
QY	1627	TTAGGAATAGGAGCGCTTTGGTGAAGTCTGTCTAGCAAGAAAGTCTACTAAAGCTTTG	1686				
Db	338	CTGGGGATCGGTGCTTTGGAGAAAGTGTGCCTTGTCTTGAAGGTGGACACTCACGCCCTG	397				

```

QY 2767 TATAATTATCAAAAGCCTATTGAGTATGTAATACATTTCACAGGCTCAGA 2819
DB 1475 TTTCGATGCCCAAAGCCTTCAGGAGCAGCAAGCTTCACAGGCTGAGAGCTCAGA 1527,

RESULT 2
; Sequence 26, Application US/09879536
; Patent NO. US20020144298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Willson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; CURRENT FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows version 3.0
; SEQ ID NO 26
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-879-536-26

```

Query Match	16.1%	Score 518.6;	DB 10;	Length 638;
Best Local Similarity	89.5%;	Pred. No. 4.6e-119;		
Matches 570;	Conservative 0;	Mismatches 64;	Indels 3;	Gaps 1;
QY 1444	GTCTGTGAAGTCTCATCAGCAGCGTCTCGATCGGAAGAGCAGCTAGAAAATGAATGATG	1503		
DB	GTACTCAAACTCATCAGCAGCGTCTACATCGTAAAAACAATTAGAGAATGAATGATG	578		
QY 1504	CGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTGCAGAAAGAG	1563		
DB	CGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTGCACAAAAGAA	518		
QY 1564	TCTAACTATATTGCTCTTTAAAGGGCTTAAATGGACAAGTCTATGTTGTAAAGATAAG	1623		
DB	TCTAATTACATCGTCTTAAAGGGCTTAAATGGACAAGTCTATGTTGTAAAGATAAG	458		
QY 1624	ACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAAGTCGATCTAAAGCT	1683		
DB	ACATTAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAAGTAGATCTAAAGCT	398		
QY 1684	TTGTATGCAACAAAGACTCTTCCAAAGAAAGCGTCTGCTCGCAAAATCAGGTGGCTCAT	1743		
DB	TTGTATGCAACAAAACACTCTTCGAATTAAGATGTCTCTTCGAATTAAGTTCGCTCAT	338		
QY 1744	GTGAAAGCGGAGGGATATCTTAGCAGAAGCCGACAATGAGTGGGTGPGCCGCTGTAC	1803		
DB	GTTAAGGCTGAGAGAGATATCTTGGCTGAAGCTGACAATGAATGGGTAGTTCGCTATAT	278		
QY 1804	TACTCTTCTCCAGCAAGGACAACCTTGTACTTTGTGATGGACTACATTCCTCGGGGGGAT	1863		
DB	TATTCATCCAAAGATAAGCAATTTATCTTTGTAATGGACTACATTCCTCGGGGTGAT	218		
QY 1864	ATGATGAGCCTATTAAATTAGAATGGGCATCTTTCCTGAAAATCTGGCAGATTTCTACATA	1923		
DB	ATGATGAGCCTATTAAATTAGAATGGGCATCTTTCGAAAAGTCTGGCAGATTTCTACATA	158		

QY 1924 GCAGAACTTACCTGTGCTGCAAGTGGGTTTATTCATAGAGAT 1980
DB 157 GCAGAACTTACCTGTGCTGCAAGTGGGTTTATTCATAGAGAT 98
QY 1981 ATTAACCTGATAACATTTTGTGATGACCTGATGGCCATATTAAATGACTGACTTTGGC 2040
DB 97 ATTAACCTGATAACATTTTGTGATGACCTGATGGCCATATTAAATGACTGACTTTGGC 38
QY 2041 TTGTGACCTGGCTTCAGATGGACATGACTCCAAGT 2077
DB 37 CTCGACCTGGCTTCAGATGGACATGACTCCAAGT 1

RESULT 3

US-09-879-536-66

; Sequence 66, Application US/09879536

; Patent No. US20020144298A1

; GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.

; APPLICANT: Steinmann, Kathleen E.

; APPLICANT: Astle, Jon H.

; APPLICANT: Burgess, Christopher C.

; APPLICANT: Bushnell, Steven E.

; APPLICANT: Carroll III, Eddie

; APPLICANT: Catino, Theodore J.

; APPLICANT: Derti, Adnan

; APPLICANT: Ford, Donna M.

; APPLICANT: Lewis, Marcia E.

; APPLICANT: Monahan, John E.

; APPLICANT: Schlegel, Robert

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; FILE REFERENCE: CCD-257 (US)

; CURRENT APPLICATION NUMBER: US/09/879,536

; PRIOR FILING DATE: 2001-09-21

; PRIOR APPLICATION NUMBER: US 60/088,801

; PRIOR FILING DATE: 1998-06-10

; NUMBER OF SEQ ID NOS: 850

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO: 66

; LENGTH: 678

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)....(678)

; OTHER INFORMATION: n = A,T,C or G

US-09-879-536-66

Query Match 11.7%; Score 376.6; DB 10; Length 678;
Best Local Similarity 82.4%; Pred. No. 8.4e-84;
Matches 500; Conservative 0; Mismatches 97; Indels 10; Gaps 6;

QY 1447 CTGAAGTCTATCAGCAGCGCTGTCATCGGAAGAAGCAGCTAGAAAATGAAATGATGCGG 1506
DB 2 CTCAATCTCATCAGCAGCGCTGTCATCGTAAAAACAATTAGAAGTGAATGATGCGG 61
QY 1507 GTTGATATATCAAGATGCCAGGATCAAAATGAGAAGATGCTTTGCCAGAAAGATCT 1566
DB 62 GTTGATATATCAAGATGCCAGGATCAAAATGAGAAGATGCTTTGCCAGAAAGATCT 121
QY 1567 AACTATATTCGTTTAAAGGGCTTAAATGGACAGTCTATCTTTGTAAGATAAAGACA 1626
DB 122 AATTACATCGGCTTTAAAGGGCTTAAATGGACAGTCTATCTTTGTAAGATAAAGACA 181
QY 1627 TTAGGAATAGGAGCGTTTGGTGAAGTCTGCTAGCAAGAAAAGTGCATATAAGCTTTG 1686
DB 182 CTAGGAATAGGAGCATTTGGTGAAGTCTGCTAGCAAGAAAAGTGCATATAAGCTTTG 241
QY 1687 TATGCAACAAAGACTCTTCCGAAGAAGACGTTCTGCTCCGAAATCAGGTGCTCATGTG 1746
DB 242 TATGCAACAAAGACTCTTCCGAAGAAGATGTTCTTCTCGAAATCAAGTCTCATGTT 301

QY 1747 AAAGCGGAGAGGATATCTTAGCAGAAAGCCGACAAATGAGTGGTGGTCCGCTGTACTAC 1806
DB 302 AAGGCTGAGAGAGATATCTCTGGCTGAAGCTGACAAATGAATGGTAGTCTGTATATAT 361
QY 1807 TCTTTCCAGGACAAGGACAACCTTTGACTTTGTGATGGAGTACATTCCTGGGGGGATATG 1866
DB 362 TCATTCCAAAGATAAGGGCCATTATTCCTT - GTAATGGCTACATTCCTNGGGGTGATATG 419
QY 1867 ATGAGCCTATTAAATAGATGGCATC-TTTCCTGAAATCTGGCAGGATTCACAT--A 1923
DB 420 AAGAGCCCATTAATTANAATGGGCATCTTTTCCAGAAAGGCTNGCACCATCTACCTTAG 479
QY 1924 GCAGAACTTACCTGTG-CAGTTGAAAAGTGTTCAT--AAAATGGGTTTATTTCATAGAGAT 1980
DB 480 CCAGAACTTACCTGTGNGCCNGTTGAAAGTGTCTCTTAAATGGGTTTAAATCTTAGAGAT 539
QY 1981 ATTAACCTGATAACATTTTGTATG--ACCGTGTGAGGCCATATTAATTTGACTGACTTTG 2038
DB 540 TTTTAACCTGGATAATATTTGANTGACCGGAAAGGCCCTTATTAATGCTTGGCTTTG 599
QY 2039 GCTTGTG 2045
DB 600 GCCTTNG 606

RESULT 4

US-09-764-868-214

; Sequence 214, Application US/09764868

; Patent No. US20020168711A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT232

; CURRENT APPLICATION NUMBER: US/09/764,868

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO: 214

; LENGTH: 676

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: SITE

; LOCATION: (628)

; OTHER INFORMATION: n equals a,t,g, or c

US-09-764-868-214

Query Match 11.6%; Score 371.8; DB 9; Length 676;
Best Local Similarity 73.5%; Pred. No. 1.3e-82;
Matches 497; Conservative 2; Mismatches 175; Indels 2; Gaps 2;

QY 1411 TTAAAGTTTCTCATGGAGCAGCAGCTAGAGAAGCTCTGTAAGTCTCATCAGCAGCGTCTG 1470
DB 1 TTAAAGTTTCTCATGGAGCAGCAGCTAGAGAAGTCTATCAAAACCTACCAGCAGAGGTT 60
QY 1471 CATCGGAAGAACAGCAGCTAGAAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAG 1530
DB 61 AACCGGAGGCTCGAGCTGGAGCAAGAAATGGCCAAAGCTGGAGTCTGTGAAGCTGAGCAG 120
QY 1531 GATCAATGAGAAAGATGCTTTGCCAGAAAGAGTCTTAACATATATTCGTTTAAAGGGCT 1590
DB 121 GAGCAGATCGGAAGATGCTCTACCAGAAAGAGTCTAATTACACAGGTTAAAGAGGGCC 180
QY 1591 AAAATGGACAAGTCTATGTTTCTTAAAGATAAAGACATAGGAATAGGAGCGTTTGGTGA 1650
DB 181 AAGATGGACAAGTCTATGTTTCTTCAAGATCAAAACCTGGGGATCGTGGCTTTGGAGAA 240
QY 1651 GTCTGCTAGCAAGAAAAGTGCATTAAGCTTTGTATGCAACAAAGACTCTTCGAAG 1710
DB 241 GTGTCCTTGTGTTGAAGGTGGACACTCACGCCCTGTACGCCATGAAGACCCCTAAGAA 300
QY 1711 AAAGAGTTTCTCTCCGAAATCAGGTGGTGCATGTGAAGCGGAGGAGGATATCTAGCA 1770

Fri Jan 17 11:17:00 2003

```

Db 301 AAGGATGCTGGAACCGAATCAGGTGGCCCGACGTCAGGCCGAGAGGACATCCTGGCC 360
Qy 1771 GAAGCCGACAAATGAGTGGGTGGTCCGCTGACTACTCTTTCCAGGACGAGGACACTTG 1830
Db 361 GAGCAGACAAATGAGTGGGTGG-CAAACTCTACTACTCTTCCAGCAGGACGAGCCTG 419
Qy 1831 TACTTTGTGAGTGGTACTACATCTCTGGGGGATATGATAGCCCTATTAATTAAGATGGC 1890
Db 420 TACTTTGTGAGTGGTACTACATCTCTGGGGGATATGATAGCCCTGCTGATCGGATGGAG 479
Qy 1891 ATCTTTCTGAAATCTGGCAGCAATCTACATAGCAGAACTTACCTGTGAGTGGTGAAGT 1950
Db 480 GTCTTCTGAGCAGCTGGCCGCKTCTACATCGCA-RACTGACTTTGGCCATTTGAGATG 538
Qy 1951 GTTCAATAAATGGTTTTATTCATAGAGATATTAACCTGATACATTTTGTGATGGCGT 2010
Db 539 GTCCACAAGATGGCTTCATCCACCGAGACATCAAGCCTGATACATTTTGATAGATCTG 598
Qy 2011 GATGCCCATATTAATGACTGACTTTGGCTTGTGCACTGGCTTCAAGTGGACATGAC 2070
Db 599 GATGTCACATTAACACTCACAGATTTGGNCTCTGCACTGGGTTCAAAGTGGACTCACAAT 658
Qy 2071 TCCAAGTACTACAGA 2086
Db 659 TTCCAATATTACCAGA 674

RESULT 5
US-09-938-842A-1861
; Sequence 1861, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPT1300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1861
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1861

Query Match 7.7%; Score 247.6; DB 9; Length 1689;
Best Local Similarity 51.5%; Pred. No. 1.5e-51;
Matches 624; Conservative 0; Mismatches 579; Indels 9; Gaps 2;

Qy 1414 AAGTTCTTCATGGAGCAGCAGCTAGAGACGTCCTGAAGTCTCATCAGCAGCGCTTGCAT 1473
Db 163 AAGCAATATATCGAAGATCATATTAATAATTCAGNAGAAGAGTCTCCAGGAAGAAAGAG 222
Qy 1474 CGGAAGAAGCAGCTAGAAATGAATGATCGGGTGGTATCTCAAGATGCCAGGAT 1533
Db 223 CGTCGAGCATCTTGGACAAACCTAGCTGATGCTGATCTGATGTTTGAAGCAAGATG 282
Qy 1534 CAATGAGAAAGATGCTTGGCAGAAAGAGTCAACTATCTGCTTAAAGGGCTAAA 1593
Db 283 GATATATTAAGAACTTTGACAAAAGGAAATGGAGTATATGCGTCTACAGACAGAAA 342
Qy 1594 ATGGACAAGTCTATCTTTGTTAAAGATAAAGACATTTAGGAATAGAGCGTTGGTGAAGTC 1653

```

RESULT 6

US-09-974-298-152

```

; Sequence 152; Application US/09974298
; Patent No. US20020156263A1
; GENERAL INFORMATION:
; APPLICANT: Chen, Hwei-Mei
; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER
; FILE REFERENCE: PA-0037 P
; CURRENT APPLICATION NUMBER: US/09/974,298
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: 60/238,331
; PRIOR FILING DATE: 2000-05-10
; NUMBER OF SEQ ID NOS: 194
; SOFTWARE: PERL Program
; SEQ ID NO 152
; LENGTH: 3583
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc_feature
; OTHER INFORMATION: Incyte ID No. US20020156263A1 347975.11
US-09-974-298-152

Query Match          7.6%; Score 244.4; DB 9; Length 3583;
Best Local Similarity 53.4%; Pred. No. 1.4e-50;
Matches 608; Conservative 0; Mismatches 491; Indels 39; Gaps 3;

Qy 1478 AGAAGCAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAA 1537
Db 1478 AGAAGCAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAA 1537

Qy 1479 AAGAGAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAA 1538
Db 1479 AAGAGAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTCAAGATGCCAGGATCAAA 1538

Qy 1538 TGAGAAAGATGCTTTGTCAGAAAGAGTCTTAAGTATATTCGCTTTAAAGGGCTAAATGG 1597
Db 1538 TGAGAAAGATGCTTTGTCAGAAAGAGTCTTAAGTATATTCGCTTTAAAGGGCTAAATGG 1597

Qy 1539 GAGATGAGCAGATCTCGGAGGAAAGAGAGTCTTCTGTTTGAAGAGAAAGAGCTTG 536
Db 1539 GAGATGAGCAGATCTCGGAGGAAAGAGAGTCTTCTGTTTGAAGAGAAAGAGCTTG 536

Qy 1598 ACAAGTCTATGTTCTGAAGATAAAGACATAGGATAGGAGCGTTTGGTGAAGTCTCTC 1657
Db 1598 ACAAGTCTATGTTCTGAAGATAAAGACATAGGATAGGAGCGTTTGGTGAAGTCTCTC 1657

Qy 1657 GATGGAAGATTTGAGTCCGTAAGAGTAAGAGTAAAGAGGAGGATTTGGTGAAGTCTCTC 596
Db 1657 GATGGAAGATTTGAGTCCGTAAGAGTAAGAGTAAAGAGGAGGATTTGGTGAAGTCTCTC 596

Qy 1658 TAGCAAGAAAGTGCATATAAGCTTTGTTATGCAACAAAGAGCTCTCGAAAGAAAGAGC 1717
Db 1658 TAGCAAGAAAGTGCATATAAGCTTTGTTATGCAACAAAGAGCTCTCGAAAGAAAGAGC 1717

Qy 1717 TTGTTCAAGAGAGATACAGGAGATGTTGATGATCAATGAAATATCTCCGTAAGAGAGATA 656
Db 1717 TTGTTCAAGAGAGATACAGGAGATGTTGATGATCAATGAAATATCTCCGTAAGAGAGATA 656

Qy 1718 TTCCTCTCGAAATCAGTGGCTCATGTGAAAGCGGAGAGGATATCTTAGCAGAAAGCG 1777
Db 1718 TTCCTCTCGAAATCAGTGGCTCATGTGAAAGCGGAGAGGATATCTTAGCAGAAAGCG 1777

Qy 1777 TCGTTGAAAAGAGAGAGGTTGGCCACATTCGTGCGGAGCGTGACATTTAGTGGAGGAG 716
Db 1777 TCGTTGAAAAGAGAGAGGTTGGCCACATTCGTGCGGAGCGTGACATTTAGTGGAGGAG 716

Qy 1778 ACAATGAGTGGTGGTGGCTGTACTACTCTTTCCAGSAGAGAGACACTGTACTTTG 1837
Db 1778 ACAATGAGTGGTGGTGGCTGTACTACTCTTTCCAGSAGAGAGACACTGTACTTTG 1837

Qy 1837 ACAGTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 776
Db 1837 ACAGTTGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 776

Qy 1838 TGATGAGTACATTTCTCGGGGGGATATGATGAGCTTATTAATAGATGGGCACTTTTC 1897
Db 1838 TGATGAGTACATTTCTCGGGGGGATATGATGAGCTTATTAATAGATGGGCACTTTTC 1897

Qy 1897 TCATGGAGTCTCTGCTGGAGGAGATGATGACCTTGTGATGAAAGAGACACTCTGA 836
Db 1897 TCATGGAGTCTCTGCTGGAGGAGATGATGACCTTGTGATGAAAGAGACACTCTCTGA 836

Qy 1898 CTGAAATCTGCGAGATTTACATAGCAGAACTTACCTGTGCACTTGAAGTGTTCATA 1957
Db 1898 CTGAAATCTGCGAGATTTACATAGCAGAACTTACCTGTGCACTTGAAGTGTTCATA 1957

Qy 1957 CAGAGAGAGAGACTCAGTTTTTATATAGCAGAAACAGTATTAAGCCATAGACTTATTACC 896
Db 1957 CAGAGAGAGAGACTCAGTTTTTATATAGCAGAAACAGTATTAAGCCATAGACTTATTACC 896

Qy 1958 AAATGGGTTTTTATCATAGAGATTTAAACCTGATAACATTTTGAATGACCGTGATGCC 2017
Db 1958 AAATGGGTTTTTATCATAGAGATTTAAACCTGATAACATTTTGAATGACCGTGATGCC 2017

Qy 1959 AACTTGGATTTATCCAGAGAGATCAAAACAGACACCTTCTTTTGGACAGCAAGGCC 956
Db 1959 AACTTGGATTTATCCAGAGAGATCAAAACAGACACCTTCTTTTGGACAGCAAGGCC 956

Qy 2017 ATATTAAATGACTGACTTTTGGCTTTGTCAGTGGCTTCAGATGGACATGACTCCAAAT 2077
Db 2017 ATATTAAATGACTGACTTTTGGCTTTGTCAGTGGCTTCAGATGGACATGACTCCAAAT 2077

Qy 2077 ACTACAGAGTGGGATCACCACCGGAGAGATAGATGATGATTTTCAATAGCAAGTGGGAG 2137
Db 2077 ACTACAGAGTGGGATCACCACCGGAGAGATAGATGATGATTTTCAATAGCAAGTGGGAG 2137

Qy 2137 TTTATGAGAAATCTGAACACAGACCTCCCGCAGTGTATTTTCCAGTCTTCCAGAAC 1066
Db 2137 TTTATGAGAAATCTGAACACAGACCTCCCGCAGTGTATTTTCCAGTCTTCCAGAAC 1066

Qy 2138 ATCCCTTCCAAATGTCGGTGTGGGAGACAGACTCAAGCCACTGGAGCGGAGAGTGTCTGCC 2197
Db 2138 ATCCCTTCCAAATGTCGGTGTGGGAGACAGACTCAAGCCACTGGAGCGGAGAGTGTCTGCC 2197

Qy 1067 -----ATGAATTCAAAAGGAAAGCAGAACTCTGGAAAAG 1101
Db 1067 -----ATGAATTCAAAAGGAAAGCAGAACTCTGGAAAAG 1101

```

```

Qy 2198 AGCACCAGCGATGCTAGCCCATTTCTGTTGGAGCTCCCAATTTATATTGACCTGAAG 2257
Db 2198 AGCACCAGCGATGCTAGCCCATTTCTGTTGGAGCTCCCAATTTATATTGACCTGAAG 2257

Qy 2258 TGCTACTGGAGACAGATATACAGAGCTGTGTGAGTGTGGAGTGTGGTGTATTTT 2317
Db 2258 TGCTACTGGAGACAGATATACAGAGCTGTGTGAGTGTGGAGTGTGGTGTATTTT 2317

Qy 2317 TGTTTATGATGAGAGCTGTGTGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAG 1220
Db 2317 TGTTTATGATGAGAGCTGTGTGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAGTGTGGAG 1220

Qy 2318 GTGAAATGTTGGTGGAGCAACCTCTCTTCTGGCACAACCCCATTTAGAAACACAAATGA 2377
Db 2318 GTGAAATGTTGGTGGAGCAACCTCTCTTCTGGCACAACCCCATTTAGAAACACAAATGA 2377

Qy 2377 ATGAGATGCTCATCGCTACCCACCTTCTGTTGAGAGACCTCAAGAGACATATAAGA 1280
Db 2377 ATGAGATGCTCATCGCTACCCACCTTCTGTTGAGAGACCTCAAGAGACATATAAGA 1280

Qy 2378 AGTTTATCATCTGGCAAACTTCTCTACACATCCCTCCCTCAAGCTAAGCTAGCTCTGAAG 2437
Db 2378 AGTTTATCATCTGGCAAACTTCTCTACACATCCCTCCCTCAAGCTAAGCTAGCTCTGAAG 2437

Qy 2437 AGGTGATGACTGGAAGAACTTTGACTTTTCTCCAGAGTTCCTCCATCTCTGAGAAG 1340
Db 2437 AGGTGATGACTGGAAGAACTTTGACTTTTCTCCAGAGTTCCTCCATCTCTGAGAAG 1340

Qy 2438 CCTCTGACCTCATTTATCAAACTGTGTGAGGACACAGAGACCGCTCGGCAAGAACGGTG 2497
Db 2438 CCTCTGACCTCATTTATCAAACTGTGTGAGGACACAGAGACCGCTCGGCAAGAACGGTG 2497

Qy 2497 CCAAGGATCTAATTTTGAAGTCTGCTGTGAATGGAACATAGAAATTTGAGCTCTCTGGAG 1400
Db 2497 CCAAGGATCTAATTTTGAAGTCTGCTGTGAATGGAACATAGAAATTTGAGCTCTCTGGAG 1400

Qy 2498 CTGATGAGATAAAGCTCATCTCATTTTAAAGACCATCGATTTCTCTAGTGTCTGAGAC 2557
Db 2498 CTGATGAGATAAAGCTCATCTCATTTTAAAGACCATCGATTTCTCTAGTGTCTGAGAC 2557

Qy 2557 TTGAGGAAATAAAAGTAACTCTTTTGAAGGCTTGA---CTGGGAACATATCAGAG 1457
Db 2557 TTGAGGAAATAAAAGTAACTCTTTTGAAGGCTTGA---CTGGGAACATATCAGAG 1457

Qy 2558 AGCAGTCTGCTTCAATACATCCCTAAATCAGCATCCCAACAGATACATCCCAATTTTGA 2615
Db 2558 AGCAGTCTGCTTCAATACATCCCTAAATCAGCATCCCAACAGATACATCCCAATTTTGA 2615

Qy 2615 AGAGACCTGCTGCAATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTGA 1515
Db 2615 AGAGACCTGCTGCAATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTGA 1515

```

```

RESULT 7
US-09-938-842A-2402
; Sequence 2402, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: SUPPRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; TITLE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPL300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2402
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2402

```

```

Query Match          6.6%; Score 212.6; DB 9; Length 1452;
Best Local Similarity 49.6%; Pred. No. 6.7e-43;
Matches 610; Conservative 0; Mismatches 609; Indels 12; Gaps 2;

Qy 1414 AAGTTCTTTCATGGAGCAGCAGCTAGAGAACGTCCTTGAAGTCTCATCAGCAGCGTCTGCAT 1473
Db 1414 AAGTTCTTTCATGGAGCAGCAGCTAGAGAACGTCCTTGAAGTCTCATCAGCAGCGTCTGCAT 1473

Qy 85 AAGCAGTACATCGAATCACTACAAAGCTCAGAACAAAGACATTTCAAGAGAGGAGAAAG 144
Db 85 AAGCAGTACATCGAATCACTACAAAGCTCAGAACAAAGACATTTCAAGAGAGGAGAAAG 144

Qy 1474 CGAAGAGACAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTCAGATGCCAGGAT 1533
Db 1474 CGAAGAGACAGCTAGAAAATGAAATGATGCGGGTTGGATTATCTCAGATGCCAGGAT 1533

Qy 145 AGACGGTGGATCTTAGAAAAGAAAGTGGCGCTTCTTGGAGTGGCCAAAAGAGGAGCAATC 204
Db 145 AGACGGTGGATCTTAGAAAAGAAAGTGGCGCTTCTTGGAGTGGCCAAAAGAGGAGCAATC 204

Qy 1534 CAATGACAAAGATGCTTTTGGCAGAAAGAGCTTAAGTATATTTCTGCTTTAAAGGCGTAAA 1593
Db 1534 CAATGACAAAGATGCTTTTGGCAGAAAGAGCTTAAGTATATTTCTGCTTTAAAGGCGTAAA 1593

Qy 205 AACATGATTAAGACCTCGAGAGAAAGAGACAGAGTTTATGAGGCTTTAAAGAGAACAG 264
Db 205 AACATGATTAAGACCTCGAGAGAAAGAGACAGAGTTTATGAGGCTTTAAAGAGAACAG 264

Qy 1594 ATGACAAAGTCTATGTTTGTAAAGATAAAGACATTTAGGAATAGGAGCGCTTTTGGTGAAGTC 1653
Db 1594 ATGACAAAGTCTATGTTTGTAAAGATAAAGACATTTAGGAATAGGAGCGCTTTTGGTGAAGTC 1653

```

Db 265 ATTAGTCGACGACCTTTGAGCTTTGACTATFCAATGGAAGAGGTCCTTTTGGTGAGGTT 324
 QY 1654 TGTCTAGCAAGAAAAGTCGATCACTAAAGCTTTGTATCAACAAAGACACTCTTCGAAAGAAA 1713
 Db 325 CGCTTATCTCGGAGAGAAAGCTCTGGAATATTTATGCCATGAAGAAGTTAAAGAAATCT 384
 QY 1714 GAGCTTCTGCTCCGAAATCAGGTGCTCATGTGAAGCGGAGAGGATATCTAGCAGAA 1773
 Db 385 GAAATGGTCATGAGAGGACAGCTGAGCATGTGAGAGAGAGGAACTCTGCTGCTGAG 444
 QY 1774 GCGGACATGAGTGGTGGTGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1833
 Db 445 GTTGAAGCCATATATTTGGAAGCTTTTACTTATTCATTTTCAAGGATCCCGAGTATCTGTAT 504
 QY 1834 TTTGTGATGAGCTACATTCCTCGGGGGGATATGATGAGCTTATTAATAGAAATGGGATC 1893
 Db 505 CTGATTATGGAATATCTCCCGGTTGATATGATGACCTTCTCATGAGGAGGACACA 564
 QY 1894 TTTCTGAAATCTGGCAGGATCTTACATAGCAGAACTTACTCTGTCAGTTGAAAGTGT 1953
 Db 565 TTACGGGAAGATGTTGCCAGATTTTATATGCTCAAGTGTCTGGCCATTTGAATCCATA 624
 QY 1954 CATAAATGGGTTTTTATCATAGAGATATTAACCTGATTAACATTTTCAATGACCGTAT 2013
 Db 625 CACAGATCAACTATATTCATAGGATATCAACCTGATTAACCTTCTTTTGGACAAAGAT 684
 QY 2014 GGCCATATTAATGACTGCTTTGGCTTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2073
 Db 685 GGGCAGATGAACTCTCGGACTTTGGCTCTGTAAGCTCTTGTATGTAGAAATTTACCT 744
 QY 2074 AAGTACTACAGAGTGGG-----GATCACCAGCGGAAGATAGATGATGATTTCACT 2124
 Db 745 TCAATTCAGGAGAAATAGGCCACCGATGATGAAACTATGTCAGAACTATGATGATTTGAT 804
 QY 2125 AACGAATGGGAGATCTTCAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2184
 Db 805 AGATGCTTCTGACACTGATACAGAGAGCTGGCGCAGTCCCGAGGAACTTCTGATTTGAT 864
 QY 2185 AGAGCTGCTGCCAGCAGCAGCTGCTAGCCATTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2244
 Db 865 ---CATTTGGCAGATGAATCGCAGAAACTAGCATTTTCACTGCTGGAACACCGGACTAT 921
 QY 2245 ATTGCACCTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2304
 Db 922 ATTGCTCTGAAGTTTGGCTGAAGAAAGGATATGGCATGGAATGTGATTTGGTGTCTATTA 981
 QY 2305 GGTGTTATTTCTGTAATGTTGGTGGGACAACTCTCTTTTCTGCTGCTGCTGCTGCTGCTGCT 2364
 Db 982 GGTCAATATGACGAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1041
 QY 2365 GAAACAAATGAAGTTATCATCTGCGCAACTTCTCTACACATCCCTCTCAAGCTAAG 2424
 Db 1042 TCACTTGCAGAAAGATGCTGCTTCAATGGAGAAACCATTTGAATTTCTGAGGATGCGAAG 1101
 QY 2425 CTGAGTCTGAAGCTCTGACCTCATTTATCAACTGTGCTGAGGACCGACAGACCGCTC 2484
 Db 1102 TTTTCATCTGAGGCAAAAGATCTTATCTGCAAGTTGCTGCAAGCTTGACCATAGGCTT 1161
 QY 2485 GGCAAGAACGCTGCTGATAGATAAAGGCTTCATCCATTTTAAAGACCATGCTTCTCT 2544
 Db 1162 GGTACTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1221
 QY 2545 AGTGATCTGAGACAGCAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2604
 Db 1222 GAAAGGCTCTATGAAATGGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1281
 QY 2605 TCCAAATTTTCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2635
 Db 1282 CAAATTTTATGAAGTTTGAAGTGAAT 1312

RESULT 8

US-09-771-161A-89
 : Sequence 89, Application US/09771161A
 : Patent No. US20020110811A1
 : GENERAL INFORMATION:
 : APPLICANT: LEVINE, et al.
 : TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
 : FILE REFERENCE: 802620-2005.1
 : CURRENT APPLICATION NUMBER: US/09/771.161A
 : CURRENT FILING DATE: 2001-01-26
 : PRIOR APPLICATION NUMBER: 09/724,676
 : PRIOR FILING DATE: 2000-11-28
 : PRIOR APPLICATION NUMBER: 136776
 : PRIOR FILING DATE: 2000-06-15
 : PRIOR APPLICATION NUMBER: 135619
 : PRIOR FILING DATE: 2000-04-12
 : NUMBER OF SEQ ID NOS: 273
 : SOFTWARE: Patentin version 3.0
 : SEQ ID NO 89
 : LENGTH: 1818
 : TYPE: DNA
 : ORGANISM: Homo sapiens
 US-09-771-161A-89
 Query Match 5.7%; Score 184; DB 10; Length 1818;
 Best Local Similarity 57.1%; Pred. No. 9.6e-36;
 Matches 354; Conservative 0; Mismatches 265; Indels 1; Gaps 1;
 QY 1478 AGAAGCAGCTAGAAAATGAATGATGCGGGTTGGATTATCTCAAGATGCCAGATCAAA 1537
 Db 729 AAAAAGATTAGAAAAGGTGATGGAAGAAGAGCCCTAAAGATGAGAGAACGACTCC 788
 QY 1538 TGAGAAAGATGCTTTGGCAGAAAAGTCTAACTATATATTCCTTTAAAGGGCTAAATGG 1597
 Db 789 GGAGATCAGACATGCTCGGAAGAACAGAGTTCTTCGTTTGAAGAGAACAGACTTG 848
 QY 1598 ACAAGTCTATGTTTGAAGATAAAGACATTAGGAATAGAGCGTTTGGTGAAGTCTGTC 1657
 Db 849 GATTGGAAGATTTTGAAGTCTCTTAAAGTAATAGCAGAGGAGCATTTGGTGAAGTACGGC 908
 QY 1658 TAGCAAGAAAAGTCGATCTAAAGCTTTGATCAACAAGACTCTTCGAAAAGAGACG 1717
 Db 909 TTGTTCAAGAAAAGATAGCGGACATGTGTATGCAATGAAATACTCCCGTAAAGCAGATA 968
 QY 1718 TTCTGCTCCGAAATCAGTGGCTTCATGTGAAAGCGGAGAGGATATCTTAGCAGAAAGCGG 1777
 Db 969 TGCTTGAAGAAAGAGCAGGTTGGCCACATTCGTCGAGCGGTGACATTTAGTGGAGCAG 1028
 QY 1778 ACAATGAGTGGTGGTCCGCTGTACTCTTTTCCAGGACAGGACACTTGTACTTTG 1837
 Db 1029 ACAGTTTGGTGGTGTGAAATTTCTATAGTTTTCAGGATAGCTTAAACCTCTACCTAA 1088
 QY 1838 TGATGACTACATTCCTCGGGGGATATGATGAGCTTATTAATAGAATGGGACTTTTC 1897
 Db 1089 TCATGAGTTCYTCCTCGGAGGACATGATGACCTTTGTATGAAAAAGACA-CTCTG 1147
 QY 1898 CTGAAAATCTGGCAGGATTTCTACATAGCAGAACTTACCTGTGCGAGTTGAAAGTTTATA 1957
 Db 1148 ACAAGAGGAGACTTCAGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTTATTCACC 1207
 QY 1958 AAATGGTTTTTATTCATAGATATTAACCTGATTAACATTTTGTATGACCGTGGTGGC 2017
 Db 1208 ACTTGGATTTCCACAGAGACATCAAAACCCAGAACCTCTTTTGGACAGCAAGGGCC 1267
 QY 2018 ATATTAAATGACTGCTTTGGCTTGTGCTTGTGCTTGTGCTGCTGCTGCTGCTGCTGCTGCT 2077
 Db 1268 ATCTGAACTTCTGACTTTGGTCTTTGCAAGGACTGAAAAAGCACATAGACAGAAAT 1327
 QY 2078 ACTACAGAGTGGGATCAC 2097
 Db 1328 TTTATAGGAATCTGAACCAC 1347

RESULT 9

US-09-771-161A-38
; Sequence 38, Application US/09771161A
; Patent No. US2002011081A1
; GENERAL INFORMATION:
; APPLICANT: LEVINE, et al.
; TITLE OF INVENTION: VARIANTS OF PROT

FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771.161A
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/724, 676
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 136776
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 135619
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 273
SOFTWARE: PatentIn version 3.0
SEQ ID NO 38
LENGTH: 1244
TYPE: DNA
ORGANISM: Homo sapiens
US-09-771-161A-38

Query Match 4.1%, Score 132.8; DB: 10; Length 1244;
Best Local Similarity 58.0%; Pred. No. 4e-23;
Matches 255; Conservative 0; Mismatches 182; Indels 3; Gaps 1;

QY 1609 TTTGTAAAGATAAAGACATTAAGTAGGAGCGCTTTGGTGAAGTCCTGTCTAGCAAGAAAA 1668
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 589 TTTGAAATAAATTAAAGTAATTTGGAAGAGGTGCTTTTGGTGAGGTGCTCTTCTCAAATG 648

QY 1669 GTCCGATCTAAAGCTTTTGTATCCACAAGAGACTCTTCGAAAGAACACGTTCTGCCTCGA 1728
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 649 AGAATACTGAACGAATTTATGCAATGAAAATCCTCAACAAGTGGGAGATGCTTGAAAGA 708

QY 1729 AATCAGGTGGCTCATGTGAACGCCGAGGAGATCTCTAGCAAGCCGACAAATGAGTGG 1788
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 709 GCAGAGACC CGCTGCTCCGAGAGAGCGCGCATGCTGTGTGAACGCGACTGCCAGTGG 768

QY 1789 GTGGTCCGCTGTACTACTCTTTCCAGACAAGACAACTTTGATTTGTGATGGACTAC 1848
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 769 ATCACCGCGCTGCACCTACGCTTTCCAGACGAGAACCCCTGTACTTAGTCATGGATTAC 828

QY 1849 ATTCTCGGGGGGATGATGATGACCTTAAATTAGAAT--GGCATCTTTCTCTGAAAT 1905
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 829 TAGTGGGTGGTGATTTACTGACCTCTCACCANAATTTGAAGACAAGTTCCGGAAGAT 888

QY 1906 CTGCCACGATTCTACATAGACAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGT 1965
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 889 ATGCGAGGTTCTACATTTGTAATGTTGGTGGCCATTGACTCCATCCACCTTCAT 948

QY 1966 TTTATCATAGAGATTAACCTGATAACATTTTGATTGACCGCTGTGCGCATATTAAA 2025
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 949 TACGTGCACAGACATTAACCTGCAATGCTCTTTGGGAGCTGAATGGTCAATATCCG 1008

QY 2026 TTGACTGACTTTGGCTTGTG 2045
||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
Db 1009 CTGGCTGACTTTGGATCATG 1028

RESULT 12
US-09-771-161A-37
Sequence 37, Application US/09771161A
Patent No. US20020110811A1
GENERAL INFORMATION:
APPLICANT: LEVINE, et al.
TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771.161A
CURRENT FILING DATE: 2001-01-26
PRIOR APPLICATION NUMBER: 09/724, 676
PRIOR FILING DATE: 2000-11-28
PRIOR APPLICATION NUMBER: 136776
PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 135619
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 273
SOFTWARE: PatentIn version 3.0
SEQ ID NO 37

Qy 1729 AATCAGTGGCTATGTGAAGCGGAGGAGGATATCTTAGCAGAAAGCGGACAAATGAGTGG 1788
 Db 417 GCAGAGACCGCGTGTCTCCGAGAGGAGCGGATGTGCTGTGAAGCGGAGTGGCAGTGG 476
 Qy 1789 GTGGTCCGCTGTACTACTCTTCCAGGACAAGGACAACCTTGTACTTTGTGATGAGTAC 1848
 Db 477 ATCACCAGCGCTGCACTACGCTTTTCAGGACGAGAACCCCTGTACTTGTATGATGATAC 536
 Qy 1849 ATTCTGGGGGATATGATGAGCTATTAATAGAAAT---GGGCATCTTTCTCTGAAAT 1905
 Db 537 TATGTGGGTGATTTACTGACCTGCTCASCANAATTTGAGACAACCTTCCGGAAGAT 596
 Qy 1906 CTGGCAGATTTACATACGAGAACTTACCTGTGCGAGTTGAAAGTGTTCATAAAATGGGT 1965
 Db 597 ATGCGAGGTTCTACATTTGGTGAATGTGCTGGCCATTGACTCCATCCATCAGCTTCAT 656
 Qy 1966 TTTATTACATAGATATTAACCTTGATAACATTTTGTATGACCTGATGGCCATATTA 2025
 Db 657 TACGKGCACAGACATTAACCTTGACAATGTCTTTTGGACGTGAATGGTCAATATCCSC 716
 Qy 2026 TTGACTGACTTTGGCTTGTG 2045
 Db 717 CTGCTGACTTTGGATCATG 736

RESULT 14
 US-09-938-842A-633
 ; Sequence 633, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; FILE REFERENCE: SAME, AND METHODS OF USE
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR APPLICATION NUMBER: US 60/264,647
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 633
 ; LENGTH: 1398
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 US-09-938-842A-633

Query Match 4.1%; Score 130.2; DB 9; Length 1398;
 Best Local Similarity 55.4%; Pred. No. 1.9e-22;
 Matches 252; Conservative 0; Mismatches 203; Indels 0; Gaps 0;

Qy 1588 GCTAAATGGAAGTCTATCTTTGTTAAAGATAAGACATTTAGGAATAGGAGCGTTTGGT 1647
 Db 379 GGTGTAGTACATATGATGATTTTGAAGTTATGAGGTTGTGGGAAAGGTGCGTTTGGG 438
 Qy 1648 GAAGTCTGTCTAGCAAGAAAAGTCTACTAAAGCTTTGTATGCAACAAAGACTCTTCCA 1707
 Db 439 AAAGTCTACAGGTGAGGAAAAGGAGACCTCTGAGATATACGCTATGAAAGTCAATGAGG 498
 Qy 1708 AAGAAAGAGCTTCTGCTCCGAAATCAGGTGCTCATGTGGAAGCGGAGGAGATATCCTA 1767
 Db 499 AAAGATCACATTATGGAGAGAAGCAATGCTGAATACATGAAGAGTGAAGCGGATATCTT 558
 Qy 1768 GCAGAGCCGACATGAGTGGTGGTCCGCTGTACTACTCTTCCAGGACAAGGACAC 1827
 Db 559 AAAAAATGTATCAACCCCATTTGTTCACTTAATTAATGATGAGTCAACCAAGTACAGG 618
 Qy 1828 TTGTACTTTGTGATGAGTACATCTTCTGGGGGGGATATGATGAGCTATTAAATAGATG 1887

Db 619 CTGTATCTTGTCTCGACTTTTATAAAGGAGGTCTATCTTTCTTCCAACCTATCAACCA 678
 Qy 1888 GGCATCTTTCTGAAATCTGGCAGGATCTACATAGCAGAACTTACCTGTGCAAGTTGAA 1947
 Db 679 GGGCTTTTCAGAGGAGCTTGGCTCGAGTGTACACTGCAGAAATCGTCTGCAAGTTTCC 738
 Qy 1948 AGTGTTCATAAAATGGGTTTATTTCATAGAGATATTAACCTGATAAAGATTTTGTATGAC 2007
 Db 739 CATCTCCATGAGAAGGATATGACAGAGATCTGAACCCGAAACATATCTCATGAT 798
 Qy 2008 CGTATGCCCATATTAATTTGACTGACTTTGGCTT 2042
 Db 799 ACAGATGCCACGCTGATGCTAACAGATTTTGGTTT 833

RESULT 15
 US-09-981-353-6
 ; Sequence 6, Application US/09981353
 ; Patent No. US20020160382A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Lasek, Amy W.
 ; APPLICANT: Jones, David A.
 ; TITLE OF INVENTION: GENES EXPRESSED IN COLON CANCER
 ; FILE REFERENCE: PA-0038 US
 ; CURRENT APPLICATION NUMBER: US/09/981,353
 ; CURRENT FILING DATE: 2001-10-11
 ; NUMBER OF SEQ ID NOS: 194
 ; SOFTWARE: PERL Program
 ; SEQ ID NO 6
 ; LENGTH: 2365
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; NAME/KEY: misc_feature
 ; OTHER INFORMATION: Incyte ID No. US20020160382A1 3819039CB1
 US-09-981-353-6

Query Match 3.7%; Score 117.4; DB 9; Length 2365;
 Best Local Similarity 56.5%; Pred. No. 3.8e-19;
 Matches 239; Conservative 0; Mismatches 181; Indels 3; Gaps 1;

Qy 1628 TAGGAATAGGAGCGTGTGTGAAGTCTCTAGCAAGAAAAGTCGATACATAAGCTTTGT 1687
 Db 371 TCGGAAGCGGAGTGTTCGAAAGTCTCTAGCAAGACACAGGACGAGAGTGTCT 430
 Qy 1688 ATGCAACAAAGACTCTTCGAAAGAAAGAGCTTCTGCTCCGAAATCAGGTGGCTCATGTGA 1747
 Db 431 ATGAGTCAAAAGTTTTCAGAGAAAGCAATCCTGAAAAGAAAGAGGAGGAGCATATTA 490
 Qy 1748 AAGCGGAGGAGGATATCTTAGCAGAAAGCGGACAAATGAG---TGGGTGTCCGCTGTACT 1804
 Db 491 TGTCCGAGCGGAATGTCTGTGAAGAAATGTCAAGACACCTTTCTGTGGGCTTCACT 550
 Qy 1805 ACTCTTTCCAGGACAAGGACAACCTGTACTTTGTGATGAGCTACATCTCTGGGGGGATA 1864
 Db 551 TCTCTTTCCAGACTCTGACAAATTTGACTTTGCTTCTAGACTACATTAATGGTGGAGT 610
 Qy 1865 TGATGAGCTTATTAATTAAGATGGGCTCTTTTCTGAAAAATCTGGCAGATTTACATAG 1924
 Db 611 TGTCTTACCATCTCCAGAGGGAACGCTGCTCTCTGGAACACGCGCTCGTCTTCTATGCTG 670
 Qy 1925 CAGAACTTACCTGTCCAGTTGAAAGTGTTCATAAAATGGGTTTATTTCATAGAGATTA 1984
 Db 671 CTGAAATAGCAGTGTCTGGCTTACCTGCAATTCCTCAACATCGTTTATAGAGACTTAA 730
 Qy 1985 AACCTGATAAATTTTGTATGACCGTGTATGGCCATTAATAATTAATTAATTAATTAATTA 2044
 Db 731 AACAGAGAATATTTTGTAGATTCACAGGACACATTTCTCTACTGACTTCGACTCT 790
 Qy 2045 GCA 2047
 Db 791 GCA 793

Fri Jan 17 11:17:00 2003

us-09-763-334-3.rnpb

Page 10

Search completed: January 16, 2003, 22:03:41
Job time : 133.161 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:55:17 ; Search time 2777.39 Seconds
(without alignments)
18735.612 Million cell updates/sec

Title: US-09-763-334-3

Perfect score: 3213

Sequence: 1 gtgcaacattcaattacg.....cccccaaaaaaaaaaaaaaa 3213

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: EST:*

1: em_estba:**
2: em_esthum:**
3: em_estin:**
4: em_estmu:**
5: em_estov:**
6: em_estpl:**
7: em_estro:**
8: em_htc:**
9: gb_est1:**
10: gb_est2:**
11: gb_htc:**
12: gb_est3:**
13: gb_est4:**
14: gb_est5:**
15: em_estfun:**
16: em_estom:**
17: gb_gss:**
18: em_gss_hum:**
19: em_gss_inv:**
20: em_gss_pln:**
21: em_gss_vrt:**
22: em_gss_fun:**
23: em_gss_mam:**
24: em_gss_mus:**
25: em_gss_other:**
26: em_gss_pro:**
27: em_gss_rod:**

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	671.4	20.9	673	10	BB557112
2	653	20.3	665	17	AZ340286
3	622.8	19.4	667	10	BB174219
4	614.4	19.1	648	13	BI155504
5	567	17.6	650	13	BI693046
6	548	17.1	600	12	BG081357

c	7	548	17.1	621	12	BG081236
c	8	508.4	15.8	694	10	AV708169
c	9	497.2	15.5	602	14	BQ201853
c	10	486.4	15.1	578	10	BE093083
c	11	481.4	15.0	496	9	AI594372
c	12	477	14.8	658	12	EG088986
c	13	468.2	14.6	473	9	AI430221
c	14	455	14.2	458	9	AI596344
c	15	454.6	14.1	657	14	BM971435
c	16	450.8	14.0	475	12	BF450950
c	17	434.4	13.5	555	12	BE723541
c	18	424	13.2	425	10	BB751204
c	19	422.4	13.1	930	14	BQ898648
c	20	416.6	13.0	637	9	AA910802
c	21	410.8	12.8	488	10	BE093084
c	22	408.8	12.7	412	14	W75595
c	23	406	12.6	737	14	BM950668
c	24	395.4	12.3	407	9	AA592233
c	25	395.4	12.3	453	10	BE093054
c	26	394.2	12.3	855	10	BE568367
c	27	385.4	12.0	394	17	AZ503459
c	28	369.2	11.5	559	10	AW642458
c	29	359	11.2	689	13	BI331257
c	30	358	11.1	858	12	BE973820
c	31	350.6	10.9	863	10	BE620135
c	32	349.6	10.9	875	13	BI664597
c	33	345	10.7	365	17	AZ421650
c	34	340.4	10.6	405	10	BE093082
c	35	339.6	10.6	957	14	BM801311
c	36	338.8	10.5	590	13	BM052863
c	37	338.6	10.5	403	10	BE093081
c	38	335	10.4	581	12	BG143806
c	39	333	10.4	457	13	BI034906
c	40	331.2	10.3	579	13	BI964175
c	41	325.4	10.1	736	12	EG538495
c	42	318	9.9	824	13	BI252737
c	43	308.8	9.6	780	12	BF301423
c	44	292	9.1	474	13	BG958580
c	45	291.8	9.1	962	17	CNS0396F

ALIGNMENTS

RESULT 1	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
LOCUS	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
DEFINITION	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
ACCESSION	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
VERSION	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
KEYWORDS	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
SOURCE	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
ORGANISM	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
REFERENCE	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
AUTHORS	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
TITLE	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
JOURNAL	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001
COMMENT	BB557112	BB557112	BB557112	673 bp	linear	EST 26-OCT-2001

On Aug 1, 2000 this sequence version replaced gi:9643478.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic
Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222

Fax: 81-45-503-9216
 E-mail: genome-resgsc.riken.go.jp,
URL: http://genome.gsc.riken.go.jp/
 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, K., Itoh
 M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new
 genes. Genome Res. 10 (10), 1617-1630 (2000)
 wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E.,
 Watanabe, K. M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura
 S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and
 Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system--384-format
 sequencing pipeline with 384 multicapillary sequencer. Genome Res.
 10 (11), 1757-1771 (2000)
 Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara
 Y., and Hayashizaki, Y.
 Computer-based methods for the mouse full-length cDNA
 encyclopedia: real-time sequence clustering for construction of a
 nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
 Kondo, S., Shingawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa
 K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and
 Hayashizaki, Y.
 Computational Analysis of Full-Length Mouse cDNAs Compared with
 Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
 please visit our web site (<http://genome.gsc.riken.go.jp/>) for
 further details.
 cDNA library was prepared and sequenced in Mouse Genome
 Encyclopedia Project of Genome Exploration Research Group in Riken
 Genomic Sciences Center and Genome Science Laboratory in RIKEN.
 Division of Experimental Animal Research in Riken contributed to
 prepare mouse tissues.

FEATURES
source

```

/dborgnism="Mus musculus"
/db_xref="taxon:10090"
/clone="E330026M13"
/clone_lib="RIKEN full-length enriched, 2 days pregnant
adult female ovary"
/sex="female"
/tissue_type="ovary"
/dev_stage="2 days pregnant adult"
/lab_host="DH10B"
/note Site_1: SalI; Site_2: BamHI; cDNA library was
prepared and sequenced in Mouse Genome Encyclopedia
Project of Genome Exploration Research Group in Riken
Genomic Sciences Center and Genome Science Laboratory in
RIKEN, Division of Experimental Animal Research in Riken
contributed to prepare mouse tissues. 1st strand cDNA was
primed with a primer [5'.
GAGAGAGACGGCGCCACTCGAGTGTTCCTTTTNN 3']. cDNA was
prepared by using trehalose thermo-activated reverse
transcriptase and subsequently enriched for full-length by the
cap-trapper. Second strand cDNA was prepared with the
primer adapter of sequence [5'
GAGAGAGATTCGAGTAATTAAATAATCCCCCCCCCCC 3']. cDNA
was cleaved with BamHI and XhoI. Vector: a modified
plasmid KS(+) after bulk excision from Lambda FLC I."
BASE COUNT      211 a   157 c   152 g   153 t
ORIGIN
Query Match          20.9%; Score 671.4; DB 10; Length 673;
Best Local Similarity 99.9%; Pred. No. 2.1e-123;
Matches 672; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
986 CCGTGAAAGCATCGCGGTCTCTGAACCAGAGCTGCAGACTGCTTTAGGCCCAACCCATC 1045
1 CCGTGAAAGCATCGCGGTCTCTGAACCAGAGCTGCAGACTGCTTTAGGCCCAACCCATC 60
1046 CTTCCTGGATGCCACAGCCAGTTCAGACTGTCAGCCTACCCCTTTTCTGAGGGTAGAC 1105
db bb CTTCCTGGATGCCACAGCCAGTTCAGACTGTCAGCCTACCCCTTTTCTGAGGGTAGAC 120

```

QY	1106	CTTCAAGTGTGCTGTGATCCCACTCTTCTGCTGAAGCTCCAAGCTATCAAGTTCACAC	1165
Db	121	CTTCAAGTGTGCTGTGATCCCACTCTTCTGCTGAAGCTCCAAGCTATCAAGTTCACAC	180
QY	1166	CGCCTTATCCAAACATCTGCTACCAAAACCCATCTGCTCCCTCCATATGAGTCAGTAA	1225
Db	181	CGCCTTATCCAAACATCTGCTACCAAAACCCATCTGCTCCCTCCATATGAGTCAGTAA	240
QY	1226	GTAAGCCCTGCAAGCATCAACAGCTAGCTTACCCTGAAGAGATGATAGTGAGAAGTGTG	1285
Db	241	GTAAGCCCTGCAAGCATCAACAGCTAGCTTACCCTGAAGAGATGATAGTGAGAAGTGTG	300
QY	1286	CGGACAGTGGTGACTCTGGGGATAAGAAAACACAGATTACAACCTTCAACCTTATCACTG	1345
Db	301	CGGACAGTGGTGACTCTGGGGATAAGAAAACACAGATTACAACCTTCAACCTTATCACTG	360
QY	1346	TTCGGAAAACAAAGAAAGATGAAGAACGAAGAGAGTCTCGGATTCAGAGTTACTCCCCAC	1405
Db	361	TTCGGAAAACAAAGAAAGATGAAGAACGAAGAGAGTCTCGGATTCAGAGTTACTCCCCAC	420
QY	1406	AGGCCTTTAAGTTCTTCATGGAGCAGCAGCTAGAGAACGCTCTGAAAGTCTCATCAGCAGC	1465
Db	421	AGGCCTTTAAGTTCTTCATGGAGCAGCAGCTAGAGAACGCTCTGAAAGTCTCATCAGCAGC	480
QY	1466	GTCTGCATCGGAAGAACGACGTAGAAAATGAAATGATCGGGTTGGATTATCTCAAGATG	1525
Db	481	GTCTGCATCGGAAGAACGACGTAGAAAATGAAATGATCGGGTTGGATTATCTCAAGATG	540
QY	1526	CCCAGGATCAAAATCAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCGTCTTAAAA	1585
Db	541	CCCAGGATCAAAATCAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCGTCTTAAAA	600
QY	1586	GGCGTAAATGGACAAAGTCTATGTTTGTAAAGATAAAGACATAGGAATAGGAGCGTTTG	1645
Db	601	GGCGTAAATGGACAAAGTCTATGTTTGTAAAGATAAAGACATAGGAATAGGAGCGTTTG	660
QY	1646	GTGAAGTCTGTCT	1658
Db	661	GTGAAGTCTGTCT	673

RESULT 2

AZ340286/c

LOCUS IM0072117F Mouse 10kb plasmid UNGC1M library Mus musculus genomic DNA

DEFINITION IM0072117F. DNA sequence.

665 bp

DNA linear

GSS 29-SEP

RESULT 2

A2340286/c

LOCUS

DEFINITION

...

ACCESSION
NUMBER

VERSION
KEYWORDS

KEYWORDS
SOURCE

SOURCE
ORGANISM

ORGANISM

REFERENCE

AUTHORS

TITLE

FORRESTER

JOURNAL
COMMENT

COMMENT

1

1

High quality sequence stop: 665.

FEATURES
source

1. 665
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="UUGCLM0072J17"
/lab_host="Male"
/sex="Male"
/lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
/note="Vector: PWD42nv; Purified genomic DNA from M. musculus C57BL/6J (male) was obtained from the Jackson Laboratory Mouse DNA Resource (http://www.jax.org/resources/documents/dnares/). The DNA was hydrodynamically sheared by repeated passage through a 0.005 inch orifice at constant velocity. The sheared DNA was blunt end-repaired with T4 DNA polymerase and T4 polynucleotide kinase. Adaptor oligonucleotides were ligated to the blunt ends in high molar excess. The adaptor DNA was purified and size-selected for a 9.5 to 10.5 kb range using preparative agarose gel electrophoresis. Vector DNA was prepared from a derivative of PWD42 (914732114/914732114/AF129072.1), a copy-number inducible derivative of plasmid R1. The vector was ligated with adaptors complementary to the insert adaptors and purified. The sheared, adaptor mouse DNA was annealed to adaptor vector DNA, and transformed into chemically-competent E. coli XL10-Gold (Stratagene) cells and selected for ampicillin resistance."

BASE COUNT 162 a 132 c 204 g 167 t
ORIGIN
Query Match 20.3%; Score 653; DB 17; Length 665;
Best Local Similarity 99.8%; Pred No. 9.7e-120;
Matches 664; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 429 GAATCCCCCTAGCCAGGCTCAGAGGCCATTAGTTCTCTTCAGTTGGTAGACAA-CCCA 487
DB 665 GAATCCCCCTAGCCAGGCTCAGAGGCCATTAGTTCTCTTCAGTTGGTAGACAA-CCCA 606
QY 488 TCATCATGCGAGTACTAGCAATTTAACTTTACACAGGGGGGACCTGGAGTTCAAGATG 547
DB 605 TCATCATGCGAGTACTAGCAATTTAACTTTACACAGGGGGGACCTGGAGTTCAAGATG 546
QY 548 GTGGTGGTTCAGTCTCATTTATCTGTCGACCAAAATGTCGCCACTGTTCTGTCGACTCGCG 607
DB 545 GTGGTGGTTCAGTCTCATTTATCTGTCGACCAAAATGTCGCCACTGTTCTGTCGACTCGCG 486
QY 608 AGCCACCACTCCATATCTCTGACCCAGCTAATGGAAGAGCCCTCTGCTTTACAAA 667
DB 485 AGCCACCACTCCATATCTCTGACCCAGCTAATGGAAGAGCCCTCTGCTTTACAAA 426
QY 668 CAGGGGCTTCTGCTCTCCACCATATTCGCAATGGAACGTTCTCTCAGTCGATGATGG 727
DB 425 CAGGGGCTTCTGCTCTCCACCATATTCGCAATGGAACGTTCTCTCAGTCGATGATGG 366
QY 728 TGCCCAACAGGAAACAGTACATGAGCTTTAATAATTAATCTCCCTGGACTGCAAA 787
DB 365 TGCCCAACAGGAAACAGTACATGAGCTTTAATAATTAATCTCCCTGGACTGCAAA 306
QY 788 CAGCTGCCCCAGTCGCTTCTGCTGCGCAGCTATCCCCAAGCGGTGGGCAATGAAA 847
DB 305 CAGCTGCCCCAGTCGCTTCTGCTGCGCAGCTATCCCCAAGCGGTGGGCAATGAAA 246
QY 848 TTCCTACATGGCAACCTACATACAGTACAGTCAAAATCTTTTAATACCCATTAGAA 907
DB 245 TTCCTACATGGCAACCTACATACAGTACAGTCAAAATCTTTTAATACCCATTAGAA 186
QY 908 GTAGAGCAAGTACATCTGCTTAATCTCAGCTTCTGCGCAGTACAGTCACTGCATCACAC 967
DB 185 GTAGAGCAAGTACATCTGCTTAATCTCAGCTTCTGCGCAGTACAGTCACTGCATCACAC 126
QY 968 CCGCTCTCTATTCAACAGCCCGTGAAAGCATGCGCGTCTCTGAAACAGAGCTGCGAGACTG 1027

Db 125 CCGCTCTCTATTCAACAGCCCGTGAAAGCATGCGCGTCTCTGAAACAGAGCTGCGAGACTG 66
QY 1028 CTTTAGCCCCAACCCATCTCTTCTTGATGCCACAGCAGTTCAGACTGTTTACGCTACCC 1087
Db 65 CTTTAGCCCCAACCCATCTCTTCTTGATGCCACAGCAGTTCAGACTGTTTACGCTACCC 6
QY 1088 CTTT 1092
Db 5 CTTT 1
RESULT 3
LOCUS BB174219
DEFINITION BB174219 RIKEN full-length enriched, adult male hypothalamus Mus musculus cDNA clone A23050F21 3' similar to AF104414 Mus musculus large tumor suppressor 1 (Lats1) mRNA, mRNA sequence.
ACCESSION BB174219
VERSION BB174219.2 GI:16269492
KEYWORDS EST.
SOURCE house mouse.
ORGANISM Mus musculus.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 667).
AUTHORS Arakawa, T., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hiramoto, K., Hori, F., Ishii, Y., Ito, M., Kawai, J., Konno, H., Kouda, M., Koya, S., Matsuyama, T., Miyazaki, A., Nomura, K., Ohno, M., Okazaki, Y., Okido, T., Saito, R., Sakai, C., Sakai, K., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Suzuki, H., Tagami, M., Tagawa, A., Takahashi, F., Takeda, Y., Tanaka, T., Toya, T., Muramatsu, M. and Hayashizaki, Y.
RIKEN Mouse ESTs (Arakawa, T., et al. 2001)
UNPUBLISHED (2001)
On Jun 29, 2000 this sequence version replaced gi:8833302.
Contact: Yoshihide Hayashizaki
Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Yokohama Institute
The Institute of Physical and Chemical Research (RIKEN)
1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan
Tel: 81-45-503-9222
Fax: 81-45-503-9216
Email: genome-res@gsc.riken.go.jp.
URL: http://genome.gsc.riken.go.jp/
Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Konno, H., Okazaki, Y., Muramatsu, M. and Hayashizaki, Y.
Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes. Genome Res. 10 (10), 1617-1630 (2000)
wagi, K., Fujiwaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsuura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y.
RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multicapillary sequencer. Genome Res. 10 (11), 1757-1771 (2000)
Konno, H., Fukunishi, Y., Shibata, K., Itoh, M., Carninci, P., Sugahara, Y. and Hayashizaki, Y.
Computer-based methods for the mouse full-length cDNA encyclopedia: real-time sequence clustering for construction of a nonredundant cDNA library. Genome Res. 11 (2), 281-289 (2001)
Kondo, S., Shinagawa, A., Saito, T., Kiyosawa, H., Yamanaka, I., Aizawa, K., Fukuda, S., Hara, A., Itoh, M., Kawai, J., Shibata, K. and Hayashizaki, Y.
Computational Analysis of Full-Length Mouse cDNAs Compared with Human Genome Sequences Mamm. Genome. 12, 673-677 (2001)
Please visit our web site (http://genome.gsc.riken.go.jp/) for further details.
cDNA library was prepared and sequenced in Mouse Genome Encyclopedia project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN Division of Experimental Animal Research in Riken contributed to

Db	542	ATCTGCTACACAAACCATCTGTCCCTCCATATGAGTCAGTAGTAAGCCCTGCAAG	601
QY	1241	ATGAACAGCCTAGCTTATCCCAAGCAAGATGATAGTGAAGAGTGCAGACAGTGTGACT	1300
Db	602	ATGACACACCTTAGCTTATCCCAAGCAAGATGATAGTGAAGAGTGCAGACAGTGTGACT	661
QY	1301	CTGGGG	1306
Db	662	CTGGGG	667
RESULT	4		
LOCUS	BT155504		
DEFINITION	602903902F1 NCL_CGAP_Mam3 Mus musculus cDNA clone IMAGE:5033425 5'	648 bp mRNA linear	EST 05-JUL-2001
ACCESSION	BT155504		
VERSION	BT155504.1	GI:14615505	
KEYWORDS	EST.		
SOURCE	house mouse.		
ORGANISM	Mus musculus		
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;		
AUTHORS	Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.		
TITLE	NIH-MGC http://mgc.nci.nih.gov/.		
JOURNAL	National Institutes of Health, Mammalian Gene Collection (MGC)		
COMMENT	Unpublished (1999)		
	Contact: Robert Strausberg, Ph.D.		
	Email: cgapbs@mail.nih.gov		
	Tissue Procurement: Iotnar Hennighausen Ph.D., Chu-Xia Deng Ph.D.		
	cDNA Library Preparation: Life Technologies, Inc.		
	cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)		
	DNA Sequencing by: Incyte Genomics, Inc.		
	Clone distribution: MGC clone distribution information can be		
	found through the I.M.A.G.E. Consortium/LLNL at:		
	http://image.llnl.gov		
	Plate: LLAM1093 row: d column: 02		
	High quality sequence stop: 648.		
	Location/Qualifiers		
	1. 648		
FEATURES	source		

```

FEATURES
source
High quality sequence stop: 648.
Location/Qualifiers
1. 648
/organism="Mus musculus"
/strain="129,C57BL/6J,FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5033425"
/clone_lib="NCI_CGAP_Mam3"
/tissue_type="tumor gross tissue"
/dev_stage="10 months"
/lab_host="DH10B"
/note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: Salt;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dr.
Library constructed by Life Technologies. Investigators
providing samples: Iothar Hennighausen/Chu-Xia Deng, NIH
Reference for transgenic model: Xu et al., Nature Genetics
22, 37-43 (1999)."
BASE COUNT 184 a 132 C 167 g 165 t
ORIGIN
Query Match 19.1%; Score 614.4; DB 13; Length 648;
Best Local Similarity 99.5%; Pred. No. 4.7e-112;
Matches 637; Conservative 0; Mismatches 1; Indels 2; Gaps 2;
QY 1643 TTGGTGAAGTCTGCTAGCAAGAAAAGTCGATACTAAAGCTTTGTATGCAACAAGACTC 1702
Db 11 TTGGTGAAGTCTGCTAGCAGNAAAGTCGATACTAAAGC-TTGTATGCGACAAGACTC 69
QY 1703 TTCGAAGAAAAGCGTTCTGCTCCGAAATCAGGTGGCTCATGTGAAAGCGGAGAGGATA 1762
Db 70 TTCGAAGAAAAGCG-TCTGTCTCCGAAATCAGGTGGCTCATCTGAACGGAGAGGATA 128
QY 1763 TCCTAGCAGAACCCGACAAATGATGGGTGGTCCGCCCTGTACTACTCTTTCCAGGACAAGG 1822
Db 129 TCCTAGCAGAACCCGACAAATGATGGGTGGTCCGCCCTGTACTACTCTTTCCAGGACAAGG 188

```


QY 1823 ACAACTTGTACTTTGTGATGACTACATTCCTGGGGGGATATGATGAGCCTTATTAATTA 1882
 |||||
 Db 189 ACAACTTGTACTTTGTGATGACTACATTCCTGGGGGGATATGATGAGCCTTATTAATTA 248
 |||||
 QY 1883 GAATGGGATCTTCTCTGAAATCTGGCAGCATTTACATAGCAGAACTTACCTGTGCAG 1942
 |||||
 Db 249 GAATGGGATCTTCTCTGAAATCTGGCAGCATTTACATAGCAGAACTTACCTGTGCAG 308
 |||||
 QY 1943 TTGAAGTGTTCATAAATGGTTTTATTATCATAGAGATATTAAACCTCATACATTTTCA 2002
 |||||
 Db 309 TTGAAGTGTTCATAAATGGTTTTATTATCATAGAGATATTAAACCTCATACATTTTCA 368
 |||||
 QY 2003 TTGACCGTGTGATGGCCATTTAAATTTGACTGACTTTGGGCTTTGTCACCTGGCTTCAGATGA 2062
 |||||
 Db 369 TTGACCGTGTGATGGCCATTTAAATTTGACTGACTTTGGGCTTTGTCACCTGGCTTCAGATGA 428
 |||||
 QY 2063 CACATGACTCCAAAGTACTACAGATGGGGATCACCCAGCGCAAGATAGCATGGATTTC 2122
 |||||
 Db 429 CACATGACTCCAAAGTACTACAGATGGGGATCACCCAGCGCAAGATAGCATGGATTTC 488
 |||||
 QY 2123 GTAAGCAATGGGGAGATCTTCCCAATTGTCGGTGTGGGACAGACTGAAGCCACTGGAGC 2182
 |||||
 Db 489 GTAAGCAATGGGGAGATCTTCCCAATTGTCGGTGTGGGACAGACTGAAGCCACTGGAGC 548
 |||||
 QY 2183 GGAGAGCTGCTGCCAGCAGCAGCGATGCTAGCCCATTTCTGCTGGTGGGACTTCCCAATT 2242
 |||||
 Db 549 GGAGAGCTGCTGCCAGCAGCAGCGATGCTAGCCCATTTCTGCTGGTGGGACTTCCCAATT 608
 |||||
 QY 2243 ATATTGCACTGAAGTGTCTACTGCGAAGGATATACACA 2282
 |||||
 Db 609 ATATTGCACTGAAGTGTCTACTGCGAAGGATATACACA 648
 |||||

RESULT 5
 BI693046
 LOCUS 603344592P1 NCI_CGAP_Mam2 Mus musculus cDNA clone IMAGE:5372231 5',
 DEFINITION 650 bp mRNA l'linear EST 18-SEP-2001
 mRNA sequence.
 ACCESSION BI693046
 VERSION BI693046.1 GI:15655675
 KEYWORDS EST.
 SOURCE house;mouse.
 ORGANISM Mus musculus.
 Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 650)
 NIH-MGC <http://mgc.nci.nih.gov/>.
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-remail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: Life Technologies, Inc.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: L14M11947 row: h column: 24
 High quality sequence start: 5
 High quality sequence stop: 643.
 Location/Qualifiers
 1. 650
 /organism="Mus musculus"
 /strain="FVB/N-3"
 /db_xref="taxon:10090"
 /clone="IMAGE:5372231"
 /clone_lib="NCI_CGAP_Mam2"
 /tissue_type="tumor, biopsy sample"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site_1: SalI;

Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
 Library constructed by Life Technologies. Investigator
 providing samples: Gilbert Smith, NIH

BASE COUNT 165 a 198 c 128 g 159 t
 ORIGIN

Query Match 17.6%; Score 567; DB 13; Length 650;
 Best Local Similarity 95.4%; Pred. No. 1.3e-102;
 Matches 629; Conservative 0; Mismatches 20; Indels 10; Gaps 4;

QY 470 CAGTGTGTAGACAACCCATCATCGAGAGTACAGAAATTTAACTTTACACAGGCG 529
 |||||
 Db 2 CTGTGAGTGTAGACACCATCATCGAGTACAGCAATT---AATTCACAGG-- 55
 |||||
 QY 530 GACCTGGAGTTTCAGAAATGCTGCTGCTGATTTATCTGTCACCAAAATGTCGCCA 589
 |||||
 Db 56 --CGACTGAGTCAAGATGCTGCTGCTGATTTATCTGTCACCAAAATGTCGCCA 113
 |||||
 QY 590 CTGTTCTGTGACTCGGAGCCACCATCTCCATATCTCTGACCCAGCAATGACAAA 649
 |||||
 Db 114 CTGG-TCTGTGACTCGGAG-CACCACCTCCATATCTCTGACCCAGCAATGACAAA 171
 |||||
 QY 650 GCCCTCTGCTTTACAAACAGGGGCTTCTGCTGCTCCACCATCATTCGCCAATGGAAG 709
 |||||
 Db 172 GCCCTCTGCTTTACAAACAGGGGCTTCTGCTGCTCCACCATCATTCGCCAATGGAAG 231
 |||||
 QY 710 TTCTCAGTGTGATGTGCTCCCAACAGAGCAAGTATACATGAGCTTTATATATTA 769
 |||||
 Db 232 TTCTCAGTGTGATGTGCTCCCAACAGAGCAAGTATACATGAGCTTTATATATTA 291
 |||||
 QY 770 ATGTCCTCTGGACTGCAAAACAGCTGGCCCGAGCTGCTCTCTGCTGCGAGTCATCCC 829
 |||||
 Db 292 ATGTCCTCTGGACTGCAAAACAGCTGGCCCGAGCTGCTCTCTGCTGCGAGTCATCCC 351
 |||||
 QY 830 CAAGCGTGGGCATGAATTCCTACATGCGAACCTTAACATACAGTGGAGTCAATTCCT 889
 |||||
 Db 352 CAAGCGTGGGCATGAATTCCTACATGCGAACCTTAACATACAGTGGAGTCAATTCCT 411
 |||||
 QY 890 TTAATAACCCATTAGGAAGTAGCAAGTCACTCTGCTAAATTCAGGCTTCGCGCACTA 949
 |||||
 Db 412 TTAATAACCCATTAGGAAGTAGCAAGTCACTCTGCTAAATTCAGGCTTCGCGCACTA 471
 |||||
 QY 950 CAGTCACCTGCCATCACCCCGCTCTTATTCACAGCCCGTGAAAGAGTCGCGCTCTGA 1009
 |||||
 Db 472 CAGTCACCTGCCATCACCCCGCTCTTATTCACAGCCCGTGAAAGAGTCGCGCTCTGA 531
 |||||
 QY 1010 AACGAGCTGCGAGTGTGTTAGCCCAACCCATCTCTTCTGGATGCGACAGGCTTC 1069
 |||||
 Db 532 AACGAGCTGCGAGTGTGTTAGCCCAACCCATCTCTTCTGGATGCGACAGGCTTC 591
 |||||
 QY 1070 AGACTGTTCAGCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCATCCCA 1128
 |||||
 Db 592 AGACTGTTCAGCTACCCCTTTTCTGAGGGTACAGCTTCAAGTGTGCTGTCATCCCA 650
 |||||

RESULT 6
 BG081357/c
 LOCUS BG081357
 DEFINITION H3064C02-5 NTA Mouse 15K cDNA Clone Set Mus musculus cDNA clone
 H3064C02 5', mRNA sequence.
 ACCESSION BG081357
 VERSION BG081357.1 GI:12563925
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus.
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 600)
 Kargul, G.J., Dudekula, D.B., Olan, Y., Lim, M.K., Jaradat, S.A., Tanaka
 T.S., Carter, M.G. and Ko, M.S.H.
 Verification and initial annotation of NIA mouse 15K cDNA clone set
 Unpublished (2001)
 Other_ESTs: H3064C02-3

unselected transcripts from extraembryonic tissue of 7.5-day mouse embryos reveals enrichment in the t-complex and under-representation on the X chromosome, 1998, Hum Mol Genet 7: 1967-1978.

```
BASE COUNT      169 a 120 c 139 g 193 t
ORIGIN
Query Match      17.1%; Score 548; DB 12; Length 621;
Best Local Similarity 99.1%; Pred. No. 7.6e-99;
Matches 551; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 2375 TGAAGGTTATCATCTGGCAAACTTCTTACACATCCCTCCTCAAGCTAAGCTAGTGCPCG 2434
Db 556 TAAAGGTTATCACTGGCAAACTTCTTACATCCCTCCTCAAGCTAAGCTAGTGCPCG 497
QY 2435 AAGCCTCTGACCTCATATCAACCTGTCTGAGGAGCCAGAACCCCTCGCCCAAGAGC 2494
Db 496 AAGCCTCTGACCTCATATCAACCTTTGTGAGGAGCCAGAACCCCTCGCCCAAGAGC 437
QY 2495 GTGCTGATGAGATAAAGGCTATCCATTTTAAAGACCATCGATTCTCTAGTATCTGA 2554
Db 436 GTGCTGATGAGATAAAGGCTATCCATTTTAAAGACCATCGATTCTCTAGTATCTGA 377
QY 2555 GACACAGTCTGCTTATACATCCCTTAAATCACGCATCCCAACAGATACATCCAAATTCG 2614
Db 376 GACACAGTCTGCTTATACATCCCTTAAATCACGCATCCCAACAGATACATCCAAATTCG 317
QY 2615 ACCCTGTTGATCCTGATTAATTTGTGGAGGATGGCAGGAGGAGGAGGAGGAGGAGG 2674
Db 316 ACCCTGTTGATCCTGATTAATTTGTGGAGGATGGCAGGAGGAGGAGGAGGAGGAGGAG 257
QY 2675 CTCTGAGGGATGGTATAAAATGGAGACCCCGAGCAGCTTCTTATGAGTTCACCT 2734
Db 256 CTCTGAGGGATGGTATAAAATGGAGACCCCGAGCAGCTTCTTATGAGTTCACCT 197
QY 2735 TTCGAGGTTTTTGTGATGACAAATGGCTACCCATATAATTAATTAATTAATTAATTAAT 2794
Db 196 TTCGAGGTTTTTGTGATGACAAATGGCTACCCATATAATTAATTAATTAATTAATTAAT 137
QY 2795 AATACATTCATTCACAGGCTCAGAACACACAGCTGTGATGAAGATGATCAACACAGCT 2854
Db 136 AATACATTCATTCACAGGCTCAGAACACACAGCTGTGATGAAGATGATCAACACAGCT 77
QY 2855 CCGATGGAACACACAGCTAGTCTATGTTTAAATAAAGTATGAGATCATTTGTAAGAT 2914
Db 76 CCGATGGAACACACAGCTAGTCTATGTTTAAATAAAGTATGAGATCATTTGTAAGAT 17
QY 2915 TTGCAAGAGGCTGAA 2930
Db 16 TTGCAAGAGGCTGAA 1

RESULT 8
AV708169 LOCUS
DEFINITION AV708169 ADC Homo sapiens cDNA clone ADCALE11 5', mRNA sequence.
ACCESSION AV708169
VERSION AV708169.1 GI:10725434
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 694)
Peng, Y., Song, H., Huang, Q., Huang, C., Gu, Y., Yang, Y., Gao, G., Xiao
, H., Xu, X., Li, N., Qian, B., Liu, F., Qu, J., Gao, X., Cheng, Z., Xu, Z.,
Zeng, L., Xu, S., Gu, W., Tu, Y., Jia, J., Fu, G., Ren, S., Zhong, M., Lu
, G., Hu, R., Chen, J., Chen, Z. and Han, Z.
Homo sapiens cDNA ADC clones
Unpublished (2000)
Contact: Zeguang Han
Chinese National Human Genome Center at Shanghai
351 Guo Shoujing Road, Zhangjiang Hi-Tech Park, Pudong, Shanghai
```

201203, P. R. China
Tel: 86-21-50801919(ex.45)
Fax: 86-21-50801922
Email: hanzgchgc.sh.cn
This clone is available at CHGC in Shanghai.

```
FEATURES
Location/Qualifiers
1..694
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="ADCALE11"
/clone_lib="ADC"
/tissue_type="Adrenal gland"
/dev_stage="Adult"
/lab_host="SOLR"
/note="Vector: pBluescript sk(-); Site_1: EcoRI; Site_2:
XhoI"
BASE COUNT      229 a 177 c 124 g 162 t 2 others
ORIGIN
Query Match      15.8%; Score 508.4; DB 10; Length 694;
Best Local Similarity 85.0%; Pred. No. 5.7e-91;
Matches 591; Conservative 0; Mismatches 102; Indels 2; Gaps 2;
QY 743 GTCATAACATGAGGCTTTATATATTAATGTCCTCGACTGCAACAGCCTGGCCCCAGT 802
Db 1 GTCATAACATGGAACCTATATAACATTAGTGTACCTGGACTGCAACAAATTTGGCCTCACT 60
QY 803 CGTCTTCTGCTCCTGGCAGTATCCCCAAGCGGTGGGATCAAAATTCCTACATGGCAAC 862
Db 61 CATCTTCTGCTCCAGCCCACTATCCCGAGCAGTGGGATGAAATCCCTACATGGCAAC 120
QY 863 CTAACATACCACTGAGGTCAAAATTTCTTTAATACCCATTAGGAAGTAGCAAGTCACT 922
Db 121 CTAACATACCACTGAGGTCAAAATTTCTTTAATACCCATTAGGAATAGCAAGTCACT 180
QY 923 CTGCTAATTTCTCAGCCTTCTGCCACTACAGTCACTGCCATCAACCCGCTCTCTATTCAAC 982
Db 181 CTGCTAATTTCTCAGCCTTCTGCCCTACAACAGTCACTGCAATTTACACAGCTCTCTATTCAAC 240
QY 983 AGCCGCTGAAAGCATGCGCGTCTCTCAACACAGAGCTGCAGACTGCTTTAGCCCCAACCC 1042
Db 241 AGCCGCTGAAAGATGTGCTGATTAAACACAGAGCTACAGACTGCTTTAGCACCTTACAC 300
QY 1043 ATCTCTTCTTGGATGCCACAGCCAGTTTCAGACTGTTTCAGCCTTACCCCTTTTCTGAGGGTA 1102
Db 301 ACCCTTCTTGGATACCACAGCCAAATTCAAACTGTTCAACCCAGTCTCTTTCTGAGGGAA 360
QY 1103 CAGCTTCAAGTGTGCTGTCTATCCCACTGTTGCTGGAAGCTCAAGCTATCAAGGTCCAC 1162
Db 361 CCGCTTCAAAATGTGATGCTGATGCCACCTGTTGCTGAAGCTCCAAACTATCAAGGACCAC 420
QY 1163 CACCGCTTATCCAAACATCTGTACACCAAAACCCATCTGTCTCTCATATGAGTCAG 1222
Db 421 CACCACTTACCCAAACATCTGTGTCACAAACCCATCTGTCTCTCATAGAGTCAA 480
QY 1223 TAACTAAGCCCTGCAAGATGAACAGCTTAGCTTACCCAAAGGAAGATGATAGTGAGAAGA 1282
Db 481 TCAGTAAGCTTAGCAAGAGGATCAGCCAAAGCTTGGCCCAAGGAAGATGAGAGTGAAAGA 540
QY 1283 GTGCGGACAGTGTGCTCTGGGATTAAGAAAGAAAGAAAGAGATTACAACTTCCACCTATCA 1342
Db 541 GTTATGAAATGTTGATAGTGGGATTAAGAAAGAAAGAAAGAGATTACAACTTCCACCTATTA 599
QY 1343 CTGTTCCGGAACCAAGAGATGAAGACGAGAGAGTCTCGGATTCAGAGTTACTCTCC 1402
Db 600 CTGTAGGGAACCAAGANAGATGAAGACGGAAGGAATCTCGTATTTCCAAAGTATTCTC 659
QY 1403 CACAGGCTTTTAAAGTCTTCTCATGGAGCAGCAGTA 1437
Db 660 CTCAGAG-CATTAAATCTTTATGGAGCCACATGTA 693

RESULT 9
```


REFERENCE
I (bases 1 to 496)

Fri Jan 17 11:17:01 2003

181 CCAAGCAATGGAGAGTGAACCCACCCACACCACTCAAGTTAGAGTGTACTCT 240
182 CCAAGCAATGGAGAGTGAACCCACCCACACCACTCAAGTTAGAGTGTACTCT 240
335 CCAAGCAATGGAGAGTGAACCCACCCACACCACTCAAGTTAGAGTGTACTCT 394
336 CCAAGCAATGGAGAGTGAACCCACCCACACCACTCAAGTTAGAGTGTACTCT 394
241 CCACCACTCCGAGAGGCCAGACCCACCTCCCGAGGACCACTCCCGCCCTCA 300
242 CCACCACTCCGAGAGGCCAGACCCACCTCCCGAGGACCACTCCCGCCCTCA 300
395 CCACCACTCCGAGAGGCCAGACCCACCTCCCGAGGACCACTCCCGCCCTCA 454
396 CCACCACTCCGAGAGGCCAGACCCACCTCCCGAGGACCACTCCCGCCCTCA 454
301 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACCTATCTCCCGA 360
302 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACCTATCTCCCGA 360
455 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACCTATCTCCCGA 514
456 TGGGAACCAAGCTCTCAGACAAAGCGCTACTCTGGGAACATGGAGTACCTATCTCCCGA 514
361 ATCTCCCTGTTCACCTGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCCT 420
362 ATCTCCCTGTTCACCTGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCCT 420
515 ATCTCCCTGTTCACCTGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCCT 574
516 ATCTCCCTGTTCACCTGGCGGTGGCAGAGGGGTACCTCCACCACTCTTACCCT 574
421 TCTCCCATGAATCCCTAGCCAGGCTCAGAGGCCATTAGTCTGTTCAGTTGGTGA 480
422 TCTCCCATGAATCCCTAGCCAGGCTCAGAGGCCATTAGTCTGTTCAGTTGGTGA 480
575 TCTCCCATGAATCCCTAGCCAGGCTCAGAGGCCATTAGTCTGTTCAGTTGGTGA 634
576 TCTCCCATGAATCCCTAGCCAGGCTCAGAGGCCATTAGTCTGTTCAGTTGGTGA 634
481 CAACCCATCATCAT 494
482 CAACCCATCATCAT 494
635 CTACCCCTCCACCAT 648
636 CTACCCCTCCACCAT 648

RESULT 13
LOCUS A1430221 473 bp mRNA linear EST 15-MAR-2000
DEFINITION me57h09.y1 Soares mouse embryo NDME13.5 14.5 Mus musculus CDNA
clone IMAGE:391649 5', mRNA sequence.

ACCESSION A1430221
VERSION A1430221.1 GI:4276057
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 473)
AUTHORS Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Steptoe, M., Theising, B., Allen, M., Bowers, Y., Person,
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter,
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Willson, R.
TITLE The WashU-NCI Mouse EST Project 1999
JOURNAL Unpublished (1999)
COMMENT Contact: Marra M/WashU-NCI Mouse EST Project 1999
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@wustl.edu
This clone is available royalty-free through LLNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
This read is a RESEQUENCE of a previously sequenced mouse clone.
This read has been verified (found to hit its original self in the
correct orientation)

Seq primer: -40RP from Gibco
High quality sequence stop: 453
POLYA-No. Location/Qualifiers
1. 473
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:391649"
/clone_lib="Soares mouse embryo NDME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site_1: Not 1; Site_2: Eco RI; 1st strand cDNA

FEATURES
source

373 GACACATGCTCCAACTACTACAGAGTNGGATCACCCACCGCAAGATGATGATTT 432
374 GACACATGCTCCAACTACTACAGAGTNGGATCACCCACCGCAAGATGATGATTT 432

2121 CAGTAAGCAATGGGAGAGTCTTCCAAATGTTCGGTGTGGGAGACACTGAAGCCACTGGA 2180
2122 CAGTAAGCAATGGGAGAGTCTTCCAAATGTTCGGTGTGGGAGACACTGAAGCCACTGGA 2180
433 CAGTAAGCAATGGGAGAGTCTTCCAAATGTTCGGTGTGGGAGACACTGAAGCCACTGGA 492
434 CAGTAAGCAATGGGAGAGTCTTCCAAATGTTCGGTGTGGGAGACACTGAAGCCACTGGA 492

2181 GCGG 2184
2182 GCGG 2184
493 GCGG 496
494 GCGG 496

RESULT 12
LOCUS BG088986 658 bp mRNA linear EST 26-JAN-2001
DEFINITION ut60g11.y1 Soares mouse NMGB_bcell Mus musculus CDNA clone
IMAGE:3332540 5', similar to TR:095835 095835 LARGE TUMOR SUPPRESSOR
1, mRNA sequence.

ACCESSION BG088986
VERSION BG088986.1 GI:12571583
KEYWORDS EST.
SOURCE house mouse.

ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 658)
AUTHORS NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
TITLE National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Tumor Gene Index

JOURNAL Unpublished (1997)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
This clone is available royalty-free through LLNL: contact the
IMAGE Consortium (info@image.llnl.gov) for further information.
MGI:1076704
Seq primer: -40RP from Gibco
High quality sequence stop: 473.
Location/Qualifiers
1. 658
/organism="Mus musculus"
/db_xref="taxon:10090"
/clone="IMAGE:3332540"
/clone_lib="Soares mouse NMGB_bcell"
/lab_host="DH10B (phage-resistant)"
/note="Organ: germinal B-cell; Vector: pT73D-Pac
(Pharmacia) with a modified polylinker; Site_1: Not 1;
Site_2: Eco RI; 1st strand cDNA was primed with a Not I -
oligo(dT) primer (5'
TGTTACCAATCTGAAGTGGGAGCGCGCGCTGTTTGTGTGTGTGTGTGTGTGTGT
T 3'). double-stranded cDNA was ligated to Eco RI
adaptors (Pharmacia), digested with Not I and cloned into
the Not I and Eco RI sites of the modified pT73 vector.
Library is normalized; constructed by Bento Soares and
M.Fatima Bonaldo."

FEATURES
source

BASE COUNT 180 a 206 c 135 t 1 others
ORIGIN

Query Match 14.8%; Score 477; DB 12; Length 658;
Best Local Similarity 97.8%; Pred. No. 1e-84;
Matches 483; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

QY 1 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGTCTTAAGAGTCTCTAGTTCCT 60
181 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGTCTTAAGAGTCTCTAGTTCCT 60

Db 155 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGTCTTAAGAGTCTCTAGTTCCT 214
156 GTGCAACATTCATTAACCGAAACAAAGCTGGAAGTCTTAAGAGTCTCTAGTTCCT 214

QY 61 CAGAGACAGCGGCCATCTCTAGGAGAAATGTGGTTTATCTGTAAGAGCCCAACTCA 120
62 CAGAGACAGCGGCCATCTCTAGGAGAAATGTGGTTTATCTGTAAGAGCCCAACTCA 120

Db 215 CAGAGACAGCGGCCATCTCTAGGAGAAATGTGGTTTATCTGTAAGAGCCCAACTCA 274
216 CAGAGACAGCGGCCATCTCTAGGAGAAATGTGGTTTATCTGTAAGAGCCCAACTCA 274

QY 121 CAGCGCGATGTAGGAGAGCTCTGTCTGATCCCGCATTCAGAGCTTCAGAGCTCAC 180
122 CAGCGCGATGTAGGAGAGCTCTGTCTGATCCCGCATTCAGAGCTTCAGAGCTCAC 180
275 CAGCGCGATGTAGGAGAGCTCTGTCTGATCCCGCATTCAGAGCTTCAGAGCTCAC 334
276 CAGCGCGATGTAGGAGAGCTCTGTCTGATCCCGCATTCAGAGCTTCAGAGCTCAC 334

was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
State Univ., from 2); double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M. Fatima Bonaldo.

BASE COUNT 117 a 142 c 101 g 113 t
ORIGIN

Query Match 14.6%; Score 468.2; DB 9; Length 473;
Best Local Similarity 99.4%; Pred. No. 6e-83;
Matches 470; Conservative 0; Mismatches 3; Indels 0; Gaps 0;

Qy 317 AGACAAACGCTACTCTCGGACATGGAGTAGCTAATCTCCGGAATCTCCCTGTGCCAC 376
Db 1 AGTCAAAACGCTACTCTCGGAACATGGAGTAGCTCGGAATCTCCGGAATCTCCCTGTGCCAC 60

Qy 377 CTGGGCGTGCAGGAGGGTACCTCCACCACCTCTTACCACCTCTCTCCCATGAATCCCC 436
Db 61 CTGGGCGTGCAGGAGGGCTACCTCCACCACCTCTTACCACCTCTCTCCCATGAATCCCC 120

Qy 437 CTAGCAGGCTCAGAGGCGCCATTAGTCTGTTCAGTTGGTAGACAACCCATCATGTC 496
Db 121 CTAGCAGGCTCAGAGGCGCCATTAGTCTGTTCAGTTGGTAGACAACCCATCATGTC 180

Qy 497 AGAGTACTAGCAAAATTTAACTTTACACGAGGCGACCTGGAGTTTCAGAAATGGTGTGTC 556
Db 181 AGAGTACTAGCAAAATTTAACTTTACACGAGGCGACCTGGAGTTTCAGAAATGGTGTGTC 240

Qy 557 AGTCTGATTTTATCTGTCGACCAAAATGTCCTCCACTGTTCTGTGACTCGGCGACCAACAC 616
Db 241 AGTCTGATTTTATCTGTCGACCAAAATGTCCTCCACTGTTCTGTGACTCGGCGACCAACAC 300

Qy 617 CTCCATATCTCTGACCCAGCTAATGGACAAGCCCTCTGCTTTTACAAAACAGGGGCTT 676
Db 301 CTCCATATCTCTGACCCAGCTAATGGACAAGCCCTCTGCTTTTACAAAACAGGGGCTT 360

Qy 677 CTGCTGCTCCACCATCATTCGCCAATGGAAGCTTCCTCAGTCGATGATGGTGGCCACA 736
Db 361 CTGCTGCTCCACCATCATTCGCCAATGGAAGCTTCCTCAGTCGATGATGGTGGCCACA 420

Qy 737 GGAACAGTCATAACATGAGCTTTATATATTATGTCCTCCCTGGACTGCAAAACA 789
Db 421 GGAACAGTCATAACATGAGCTTTATATATTATGTCCTCCCTGGACTGCAAAACA 473

RESULT 14
AI596344/c 458 bp mRNA linear EST 21-APR-1999
LOCUS
DEFINITION
me57h09.x1 Soares mouse embryo NBME13.5 14.5 Mus musculus cDNA
clone IMAGE:391649 3' similar to TR:Q24096 Q24096 LATS. [1] ;, mRNA
sequence.
ACCESSION
VERSION
KEYWORDS
SOURCE
EST.
AI596344
AI596344.1 GI:4605392
house mouse.
ORGANISM
Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 458)
REFERENCE
AUTHORS
Marra, M., Hillier, L., Kucaba, T., Martin, J., Beck, C., Wylie, T.,
Underwood, K., Stepec, M., Theising, B., Allen, M., Bowers, Y., Person,
B., Swaller, T., Gibbons, M., Pape, D., Harvey, N., Schurk, R., Ritter,
E., Kohn, S., Shin, T., Jackson, Y., Cardenas, M., McCann, R.,
Waterston, R., and Wilson, R.
TITLE
The Washu-NCI Mouse EST Project 1999
JOURNAL
Unpublished (1999)
COMMENT
Contact: Marra M/Washu-NCI Mouse EST Project 1999
Washington University School of Medicine

4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108, USA
Tel: 314 286 1800
Fax: 314 286 1810
Email: mouseest@watson.wustl.edu
This clone is available royalty-free through LNL; contact the
IMAGE Consortium (info@image.lnl.gov) for further information.
This clone was previously sequenced on the 5' end only, this new
data is from the 3' end
High quality sequence stop: 449.
FEATURES
Location/Qualifiers
1..458
/organism="Mus musculus"
/strain="C57BL/6J"
/db_xref="taxon:10090"
/clone="IMAGE:391649"
/clone_lib="Soares mouse embryo NBME13.5 14.5"
/sex="unknown"
/tissue_type="embryo"
/dev_stage="13.5-14.5dpc total fetus"
/lab_host="DH10B"
/note="Vector: pT73D-Pac (Pharmacia) with a modified
polylinker; Site 1: Not I; Site 2: Eco RI; 1st strand cDNA
was primed with a Not I - oligo(dT) primer [5'
TGTTACCAATCTGAAGTGGAGCGCGCGGAAATTTTTTTTTTTTTTTTTT
T 3'], on equal amounts of mRNA from 2 13.5dpc and 2
14.5dpc embryos (total RNA provided by Minoru Ko, Wayne
State Univ., from 2); double-stranded cDNA was ligated to
Eco RI adaptors (Pharmacia), digested with Not I and
cloned into the Not I and Eco RI sites of the modified
pT73 vector. Library went through one round of
normalization, and was constructed by Bento Soares and
M. Fatima Bonaldo."

BASE COUNT 98 a 98 c 119 g 143 t
ORIGIN

Query Match 14.2%; Score 455; DB 9; Length 458;
Best Local Similarity 100.0%; Pred. No. 2.5e-80;
Matches 455; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1039 ACCCATCTCTTCTGGATGCCACAGCCAGTTCAGACTGTTCAGCTTACCCCTTCTCTGAG 1098
Db 458 ACCCATCTCTTCTGGATGCCACAGCCAGTTCAGACTGTTCAGCTTACCCCTTCTCTGAG 399

Qy 1099 GGTACAGCTTCAAGTGTGCCTGTATCCCACTGTGTGCTGAAGCTCCCAAGCTATCAAGT 1158
Db 398 GGTACAGCTTCAAGTGTGCCTGTATCCCACTGTGTGCTGAAGCTCCCAAGCTATCAAGT 339

Qy 1159 CCACCACCGCTTATCCAAAACATCTGTACACCAAAACCCATCTGTCTCCCTCATATGAG 1218
Db 338 CCACCACCGCTTATCCAAAACATCTGTACACCAAAACCCATCTGTCTCCCTCATATGAG 279

Qy 1219 TCAGTAAGTAAAGCTTCAAGATGAACAGCTTACCCAGAGAGATGATGAG 1278
Db 278 TCAGTAAGTAAAGCTTCAAGATGAACAGCTTACCCAGAGAGATGATGAG 219

Qy 1279 AAGAGTGGGACAGTGTGACTCTGGGATAAAGAAAACAGATTACAACCTTCACT 1338
Db 218 AAGAGTGGGACAGTGTGACTCTGGGATAAAGAAAACAGATTACAACCTTCACT 159

Qy 1339 ATCACTGTTCGAAAAACAAGAAAGATGAAGAACGAGAGTCTCGGATTTCAGAGTTAC 1398
Db 158 ATCACTGTTCGAAAAACAAGAAAGATGAAGAACGAGAGTCTCGGATTTCAGAGTTAC 99

Qy 1399 TCCCCACAGGCTTTAAGTCTTTCATGAGCAGCAGCTAGAGAACGCTCCTGAAGTCTCAT 1458
Db 98 TCCCCACAGGCTTTAAGTCTTTCATGAGCAGCAGCTAGAGAACGCTCCTGAAGTCTCAT 39

Qy 1459 CAGCAGGCTCTGCATCGGAAGAACGAGCTAGAAAA 1493
Db 38 CAGCAGGCTCTGCATCGGAAGAACGAGCTAGAAAA 4

RESULT 15

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 5238.57 Seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagagcaccgccgaagt.....aagagcactattttggggg 3155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*

1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_on:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vl:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vl:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB	ID	Description
1	3155	100.0	3155	6	AR201457	AR201457 Sequence
2	3097.4	98.2	3460	10	AB023958	AB023958 Mus muscu
3	1749.6	55.5	5486	6	E38226	E38226 Human tumor
4	1748	55.4	5486	6	E38227	E38227 Human tumor
5	1746.4	55.4	4098	9	AF207547	AF207547 Homo sapi
6	1717.2	54.4	3412	9	AB028019	AB028019 Homo sapi
7	812.6	25.8	3213	6	AR201456	AR201456 Sequence
8	812.6	25.8	3213	10	AF104414	AF104414 Mus muscu
9	760	24.1	3984	6	AR201455	AR201455 Sequence
10	760	24.1	4241	9	AF164041	AF164041 Homo sapi
11	760	24.1	4307	9	AF104413	AF104413 Homo sapi
12	747.8	23.7	2442	6	E24613	E24613 warts prote
13	742.8	23.5	57393	2	AC094779	AC094779 Rattus no
14	676.6	21.4	1374	6	E24614	E24614 warts prote
15	566.2	17.9	5360	3	DROWARTS	L39837 Drosophila
16	564.6	17.9	5720	3	DMU296087	U29608 Drosophila
17	564.6	17.9	5720	6	AR201454	AR201454 Sequence
18	495.8	15.7	183597	9	AL356285	AL356285 Human DNA
19	460.4	14.6	74757	2	AC118930	AC118930 Mus muscu
20	366.6	11.6	16490	2	AC102712	AC102712 Mus muscu
21	356.8	11.3	181750	2	AC098623	AC098623 Rattus no
22	352.4	11.2	80146	9	AL583963	AL583963 Human DNA
23	311.6	9.9	335259	2	AC015834	AC015834 Homo sapi
24	287.2	9.1	2305	8	NCCOT1	X97657 N crassa mR
25	286.2	9.1	2231	8	AF041843	AF041843 Ustilago
26	258.8	8.2	2160	6	AR170896	AR170896 Sequence
27	258.4	8.2	1745	3	AF275634	AF275634 Caenorhab
28	251	8.0	215989	2	AC102227	AC102227 Mus muscu
29	248.2	7.9	28133	2	AC012975	AC012975 Drosophill
30	248.2	7.9	163741	3	AC007821	AC007821 Drosophill
31	247.4	7.8	2126	3	AF238490	AF238490 Drosophill
32	243.6	7.7	1235	3	CEPROKINX	Z34989 C.elegans m
33	243.4	7.7	236109	3	AE003775	AE003775 Drosophill
34	241.8	7.7	2265	3	AY051880	AY051880 Drosophill
35	240	7.6	245067	2	AC087136	AC087136 Mus muscu
36	237.8	7.5	2100	3	DMPROKINX	Z35103 D.melanogas
37	237.8	7.5	2114	3	AF239171	AF239171 Drosophill
38	237	7.5	4725	9	BC028603	BC028603 Homo sapi
39	237	7.5	4983	6	AX056364	AX056364 Sequence
40	237	7.5	5181	9	AB023182	AB023182 Homo sapi
41	233.8	7.4	1935	6	AR139101	AR139101 Sequence
42	231.6	7.3	1579	8	AY059113	AY059113 Arabidops
43	231.6	7.3	1894	8	AY086014	AY086014 Arabidops
44	231.6	7.3	1970	8	AY034932	AY034932 Arabidops
45	231.4	7.3	2422	8	AY128279	AY128279 Arabidops

ALIGNMENTS

RESULT 1
AR201457
LOCUS AR201457 3155 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 7 from patent US 6359193.
ACCESSION AR201457
VERSION AR201457.1 GI:20252345
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3155)
AUTHORS Xu,T., Tao,W., Wang,W., Zhang,S. and Yu,W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 7 19-MAR-2002;
FEATURES Location/Qualifiers

source	1. .3155	751 a	924 c	894 g	586 t	0;
BASE COUNT	100.0%;	Score 3155;	DB 6;	Length 3155;		
ORIGIN	Best Local Similarity 100.0%;	Pred. No. 0;	Mismatches 0;	Indels 0;	Gaps 0;	
	Matches 3155;	Conservative 0;				
Qy	1	ATGAGACCCACCCGAAAGTTTGGACCTTATCAAAAAGCTCTCAGGGAATCCGATATTC	60			
Db	1	ATGAGACCCACCCGAAAGTTTGGACCTTATCAAAAAGCTCTCAGGGAATCCGATATTC	60			
Qy	61	CTCTGCTCTTTTGCACACGAGTACGACCTTCGCGACCTCAGAGGTGAACCGGACAGATG	120			
Db	61	CTCTGCTCTTTTGCACACGAGTACGACCTTCGCGACCTCAGAGGTGAACCGGACAGATG	120			
Qy	121	CTTCAGGAGTTGGTGAATCGGGATGTGACACAGAGATGGCTGGCAGAGCGCTCACGAG	180			
Db	121	CTTCAGGAGTTGGTGAATCGGGATGTGACACAGAGATGGCTGGCAGAGCGCTCACGAG	180			
Qy	181	ACGGGACGAGTACGAGTTCGAGTTCGAGTACATCAGTACAGTACGAGTACGAGTACGAGT	240			
Db	181	ACGGGACGAGTACGAGTTCGAGTTCGAGTACATCAGTACAGTACGAGTACGAGTACGAGT	240			
Qy	241	CCCAGGATGAGCAGATTTGCGAGTTCATCAAGCAGACCTCCCGAGGAAAGGCGCTGGCG	300			
Db	241	CCCAGGATGAGCAGATTTGCGAGTTCATCAAGCAGACCTCCCGAGGAAAGGCGCTGGCG	300			
Qy	301	TCCACCCCGGTGACTCGCGGCGCCAGTTTCAGGGCCACAGGGGAAGCACTCCCATCTAC	360			
Db	301	TCCACCCCGGTGACTCGCGGCGCCAGTTTCAGGGCCACAGGGGAAGCACTCCCATCTAC	360			
Qy	361	CACACGCTGGGTGTCGAACTACGAGGGCCCGCCGCTGAGGAGAGTCCCGCGCAA	420			
Db	361	CACACGCTGGGTGTCGAACTACGAGGGCCCGCCGCTGAGGAGAGTCCCGCGCAA	420			
Qy	421	TATTTAGACTTTCTTCTTCCCTGGAGCGGAGCGGCGACCTGCGGCTGCGAGCTCACAG	480			
Db	421	TATTTAGACTTTCTTCTTCCCTGGAGCGGAGCGGCGACCTGCGGCTGCGAGCTCACAG	480			
Qy	481	CATCTCCCAAGGTTACAGCAGCAGTACAGCAAGTTCGCGCTGCGGCTGCGGCTCACAG	540			
Db	481	CATCTCCCAAGGTTACAGCAGCAGTACAGCAAGTTCGCGCTGCGGCTGCGGCTCACAG	540			
Qy	541	TATGCTCGTGTCTACTACTACGAGCAGTCTGGGTATGGGTGCGAGCTTCCTCC	600			
Db	541	TATGCTCGTGTCTACTACTACGAGCAGTCTGGGTATGGGTGCGAGCTTCCTCC	600			
Qy	601	TTCCAGAAACAGACCCACAGATGCTTATTCAGCATGGCCAGGCGCCAGGCTGGGCT	660			
Db	601	TTCCAGAAACAGACCCACAGATGCTTATTCAGCATGGCCAGGCGCCAGGCTGGGCT	660			
Qy	661	CCCGCCGCTCAGCTTTCTTCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720			
Db	661	CCCGCCGCTCAGCTTTCTTCCCTGCGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	720			
Qy	721	GGGGTACCCACCTCGGGGCGCCACCATATACATGTGTTGGCACCCGCGGCTCCACGTTT	780			
Db	721	GGGGTACCCACCTCGGGGCGCCACCATATACATGTGTTGGCACCCGCGGCTCCACGTTT	780			
Qy	781	ACTGGGAAAGCTCTGCACAGCTGTGCTGGCAGCTGCTGGCAGCTGCTGGCAGCTGCT	840			
Db	781	ACTGGGAAAGCTCTGCACAGCTGTGCTGGCAGCTGCTGGCAGCTGCTGGCAGCTGCT	840			
Qy	841	TTGTAGGAGCTGGGTCCACAGCTGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT	900			
Db	841	TTGTAGGAGCTGGGTCCACAGCTGCTGGTCTGCTGCTGCTGCTGCTGCTGCTGCT	900			
Qy	901	CTGCAGAACGAGGTCTAGAACGCTCGGGCGCGCATGTGGCTTTTCGGGCTGGCCCCAGC	960			
Db	901	CTGCAGAACGAGGTCTAGAACGCTCGGGCGCGCATGTGGCTTTTCGGGCTGGCCCCAGC	960			

Qy	961	AGGACCAACTCTTCAACAACCCACAACTGAGCCCTCACTCCCGCCGCCAACACAGGTC	1020
Db	961	AGGACCAACTCTTCAACAACCCACAACTGAGCCCTCACTCCCGCCGCCAACACAGGTC	1020
Qy	1021	ACGGCGTGACGGCGCACACATCTTCACTCTGTAAGAGCGTGGTGTGGGGCC	1080
Db	1021	ACGGCGTGACGGCGCACACATCTTCACTCTGTAAGAGCGTGGTGTGGGGCC	1080
Qy	1081	GAGCCCCACACAGCGTGGGGCCCTCGCACCCCGCTGCGGTGGCTGGCCACACACCT	1140
Db	1081	GAGCCCCACACAGCGTGGGGCCCTCGCACCCCGCTGCGGTGGCTGGCCACACACCT	1140
Qy	1141	GCCACTGAGAGCGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1200
Db	1141	GCCACTGAGAGCGCTGGAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG	1200
Qy	1201	TATGGCGGCTCCGAGCGGAGGTCGCCACCGCTCCCTATCCAAAGCACTTGTGTGCCC	1260
Db	1201	TATGGCGGCTCCGAGCGGAGGTCGCCACCGCTCCCTATCCAAAGCACTTGTGTGCCC	1260
Qy	1261	AGTAAGTCTGAGCAGTACAGCGTGGACCTGGACAGCGCTGTGCACAGCTGTGCAGAGT	1320
Db	1261	AGTAAGTCTGAGCAGTACAGCGTGGACCTGGACAGCGCTGTGCACAGCTGTGCAGAGT	1320
Qy	1321	CTCGAGGGGCGCTACTGATCTAGACGGGAGTGACAAAGCACTTGTGCACAGCTGTGCAGAGT	1380
Db	1321	CTCGAGGGGCGCTACTGATCTAGACGGGAGTGACAAAGCACTTGTGCACAGCTGTGCAGAGT	1380
Qy	1381	AAAGCTGGGAGAGCAAAAAGCAGATTACAGCTCCCGGTGCTCTCGCAGAAATAGC	1440
Db	1381	AAAGCTGGGAGAGCAAAAAGCAGATTACAGCTCCCGGTGCTCTCGCAGAAATAGC	1440
Qy	1441	AGAGTGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTATGCTTCAAATTC	1500
Db	1441	AGAGTGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTATGCTTCAAATTC	1500
Qy	1501	TTATGGAGCAACACGCTGGAGAGTTCATCAAAACCTTACCAGCAGAGGTCACCGGAGG	1560
Db	1501	TTATGGAGCAACACGCTGGAGAGTTCATCAAAACCTTACCAGCAGAGGTCACCGGAGG	1560
Qy	1561	CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTGAGGCGCAGAGGAGGATG	1620
Db	1561	CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTGAGGCGCAGAGGAGGATG	1620
Qy	1621	AGNAGATCTCTTACCAGAGGAGTCTAACTACAACCGCTGAAGAGGCGCAAGATGGAC	1680
Db	1621	AGNAGATCTCTTACCAGAGGAGTCTAACTACAACCGCTGAAGAGGCGCAAGATGGAC	1680
Qy	1681	AAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGCTTGGGGAAGTGTGCTC	1740
Db	1681	AAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGCTTGGGGAAGTGTGCTC	1740
Qy	1741	GCTTGAAGCTGGACACTCAGCTCTAGCCATGAGACTCTCAGAGAGAGGATGTC	1800
Db	1741	GCTTGAAGCTGGACACTCAGCTCTAGCCATGAGACTCTCAGAGAGAGGATGTC	1800
Qy	1801	CTGAACCGGATCAAGTGGCCCATGTCAAGCTCAGAGGAGACATCTCGGTGGAAGCAGAC	1860
Db	1801	CTGAACCGGATCAAGTGGCCCATGTCAAGCTCAGAGGAGACATCTCGGTGGAAGCAGAC	1860
Qy	1861	AATGAGTGGGTGTCAAACTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1920
Db	1861	AATGAGTGGGTGTCAAACTCTACTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	1920
Qy	1921	ATGAGCTACATACAGGGGGGATATGATGAGCTGTGATCAGATGGAGTCTTCCCT	1980
Db	1921	ATGAGCTACATACAGGGGGGATATGATGAGCTGTGATCAGATGGAGTCTTCCCT	1980
Qy	1981	GAGCACCTGGCGGCTTCTACATTTGACAGTGTGACCTTGGCCATTGAAAGTCTCCACAAG	2040
Db	1981	GAGCACCTGGCGGCTTCTACATTTGACAGTGTGACCTTGGCCATTGAAAGTCTCCACAAG	2040
Qy	2041	ATGGCTTTTATCCACCGGGACATCAAGCCTGACAACTACTCATCGACCTGGATGTCAT	2100

Db 2041 ATGGGCTTATCCACGGGACATCAAGCTGACAACTACTCATCGACCTGGATGTCAT 2100
QY 2101 ATTAAGCTACAGATTTTGGCTCTGCACCTGATTCAGTGGACTCACAAATTCAGATAC 2160
Db 2101 ATTAAGCTACAGATTTTGGCTCTGCACCTGATTCAGTGGACTCACAAATTCAGATAC 2160
QY 2161 TACCAGAAAGGGAACACATGAGACAGGACGATGAGCCGGTGACCTCTGGGACGAT 2220
Db 2161 TACCAGAAAGGGAACACATGAGACAGGACGATGAGCCGGTGACCTCTGGGACGAT 2220
QY 2221 GTTTCACACTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGGCGCAGAGCAG 2280
Db 2221 GTTTCACACTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGGCGCAGAGCAG 2280
QY 2281 CACCAGAGTGGCTGGCACAATTTCTTGTGCGGACACCAAAATACATCCTCGGAGGTG 2340
Db 2281 CACCAGAGTGGCTGGCACAATTTCTTGTGCGGACACCAAAATACATCCTCGGAGGTG 2340
QY 2341 CTTCTCCGCAAGGGGTACACGAGCTCTGTGCTGAGGCTGGTGTGATTCCTTT 2400
Db 2341 CTTCTCCGCAAGGGGTACACGAGCTCTGTGCTGAGGCTGGTGTGATTCCTTT 2400
QY 2401 GAGATGCTGGTGGGAGCGGCTTTCTTGTGGCCCCCACCACACAGAGCAGCTGAAG 2460
Db 2401 GAGATGCTGGTGGGAGCGGCTTTCTTGTGGCCCCCACCACACAGAGCAGCTGAAG 2460
QY 2461 GTGATCACTGGAGAGCAGCTGCATATCCCTACGAGCTGAGCTCAGGCTGAGGCC 2520
Db 2461 GTGATCACTGGAGAGCAGCTGCATATCCCTACGAGCTGAGGCTCAGGCTGAGGCC 2520
QY 2521 CGAGAGCTCATCAGCAAGCTGTGCTGCGCGGCTGACTGCGGCTGGCAGGATGGGCA 2580
Db 2521 CGAGAGCTCATCAGCAAGCTGTGCTGCGCGGCTGACTGCGGCTGGCAGGATGGGCA 2580
QY 2581 GATGACCTCAAGGACACACCCGTTCTTCAACACATCGACTTTTCCCGTGACATCCGAAG 2640
Db 2581 GATGACCTCAAGGACACACCCGTTCTTCAACACATCGACTTTTCCCGTGACATCCGAAG 2640
QY 2641 CAGGCTGACACCTAGCTCCCAACATCAGCAGCCACCCATGAGACCTCCAAATTTGACCG 2700
Db 2641 CAGGCTGACACCTAGCTCCCAACATCAGCAGCCACCCATGAGACCTCCAAATTTGACCG 2700
QY 2701 GTGGATGAAGAAAGCCCTTGGCAGAGGCGCAGGAGAGCAGCCGCTGGGACAG 2760
Db 2701 GTGGATGAAGAAAGCCCTTGGCAGAGGCGCAGGAGAGCAGCCGCTGGGACAG 2760
QY 2761 CTGGGCTCCCCAGCAGCAAGCATCCAGAGCAGCCTTCTATGAGTTACCTTCCGAG 2820
Db 2761 CTGGGCTCCCCAGCAGCAAGCATCCAGAGCAGCCTTCTATGAGTTACCTTCCGAG 2820
QY 2821 TTCTTCGATGACAGCGCTATCCCTTCCGCTGCGGAGCCCTCAGAGCCCGCAGAGT 2880
Db 2821 TTCTTCGATGACAGCGCTATCCCTTCCGCTGCGGAGCCCTCAGAGCCCGCAGAGT 2880
QY 2881 GCAGACCCAGGGGATGCGGACTTGAAGGTGCGGCGGAGGCTGCGAGCGGTGTACGTG 2940
Db 2881 GCAGACCCAGGGGATGCGGACTTGAAGGTGCGGCGGAGGCTGCGAGCGGTGTACGTG 2940
QY 2941 TAAGCCTCAGTTAACCACAACTCGAGGAAACCCAAATGAGATTTCTTTTCAGAGACAA 3000
Db 2941 TAAGCCTCAGTTAACCACAACTCGAGGAAACCCAAATGAGATTTCTTTTCAGAGACAA 3000
QY 3001 ACTCAAGCTTAGGATCCTTCAATTTTGTGCTGTAATGCGCAACAGAGATCAAC 3060
Db 3001 ACTCAAGCTTAGGATCCTTCAATTTTGTGCTGTAATGCGCAACAGAGATCAAC 3060
QY 3061 ATGATTTCAATTAGCCCTCTCAGGACCTTCACTGATTAACACAGTATTTTAAAAA 3120
Db 3061 ATGATTTCAATTAGCCCTCTCAGGACCTTCACTGATTAACACAGTATTTTAAAAA 3120
QY 3121 TTAGTACAGTGGAAAGACACTTATTTTGGGG 3155
|||||

Db 3121 TTAGTACAGTGGAAAGACACTTATTTTGGGG 3155
RESULT 2
LOCUS AB023958 3460 bp mRNA linear ROD 09-MAR-2000
DEFINITION Mus musculus MmLATS2 mRNA for warts/lats-like kinase, complete cds.
ACCESSION AB023958
VERSION AB023958.1 GI:7212785
KEYWORDS MmLATS2; warts/lats-like kinase;
SOURCE Mus musculus cDNA to mRNA.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (sites)
AUTHORS Yabuta,N., Fujii,T., Copeland,N.G., Gilbert,D.J., Jenkins,N.A.,
Nishiguchi,H., Endo,Y., Toji,S., Tanaka,H., Nishimune,Y. and
Nojima,H.
TITLE Structure, expression, and chromosome mapping of LATS2, a mammalian
homologue of the Drosophila tumor suppressor gene lats/warts
JOURNAL Genomics 63 (2), 263-270 (2000)
MEDLINE 20139436
REFERENCE 2 (bases 1 to 3460)
AUTHORS Nojima,H. and Fujii,T.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1999) Hiroshi Nojima, Research Institute for
Microbial Diseases, Osaka University, Department of Molecular
Genetics; Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
(E-mail:hnojima@biken.osaka-u.ac.jp, Tel:81-6-6875-3980,
Fax:81-6-6875-5192)
COMMENT Sequence updated (06-Jan-2000).
FEATURES
Source 1..3460
/organism="Mus musculus"
/db_xref="taxon:10090"
gene 1..3460
/gene="MmLATS2"
CDS 134..3262
/gene="MmLATS2"
/note="WALK"
/codon_start=1
/product="warts/lats-like kinase"
/protein_id="BAA92380.1"
/db_xref="GI:7212786"
/translation="MRPTFTATTYSNGSRORLQREGLKQPSKASTQGLLVGPNSD
TSLDAKVLGSKDASRQOQMRATPKPGYKALREIRYSLLPFANESGYTSAAEVNRQM
LOELVNAAGCDQEMAGRALQKTSRISIEAALYISKMGYLDPRNEQIVRIKQTSFGK
LAPTPVTRPSEGTGEALPSYHRLGANYEGPALEMPQYLDLFLPGAGAGTGA
QAQHPKGYSTAVEPSAHFPCTHYGRGHLLSEQPGYGVQSSSFONKTPPDAYSMA
KAQGGPSALTFPAHAGLYTASHHKPAATPPGAHPLHVLGTGPTFTGESSAQAVLAP
SRNSLNADLYELGSTVPSAAPLARRDSLOKGLEASRPHVAFRAGPSRTSNPNQP
EPSLPAPNTVTAAHLLHPKSVRLPEPTAVGSPHAWAAPTAPATESLETK
EGSAGPHLDVYDGGSEKRRCPYPYKHLPSKSEQSYVDLSLCTSVQOSLRGTE
QDRDCKSHKAGDKAGDKKQIQITSPVVRKNSRDEKRESIKSYSPAFKFMEO
HVENVITYQOKSVRRLOEVAKEAGCEAOEMKILYKESNVNRLKRAKMDKS
MFVKITLIGAFGEVCLACKLDTHALVAMKTLRKKDVLNRQVAVHVAERDILAEAD
NEWVVKLYSFQDRKSLYFYMDYIFGGDMSSLIRMEVFPPEHLARFYAELTLAESV
HKMGFIHRDIKPNILDLIDGLCTGFRTWTHNSKIYKGNHMRDSDMEPGD
LMDVSNCRGDRKLTLEQAKQORCLASHLVGTPNYIAPEVLLRKYTOLCDWNS
VGVLTFEMLVGQPPFLAPTPTQTKVINWSTLHPTQVRLSAEARDLITKLCAD
CRLRGDADDLKAHPFNTIDFSDIRKQAPAPYPTISHPMDISNFDVDESPWHEA
SGSAKAWDTLASPSKHPHAHYEFTFRFDDNDYFPCRCPKPSEPAESADPADL
EGAAEGCOPVIV"
polyA_site 3440
/gene="MmLATS2"
BASE COUNT 861 a 992 c 963 g 644 t
ORIGIN
Query Match 98.2%; Score 3097.4; DB 10; Length 3460;
Best Local Similarity 99.2%; Pred. No. 0;
Matches 3113; Conservative 0; Mismatches 26; Indels 0; Gaps 0;
QY 1 ATGAGAGCCACCCCGAAGTTTGACCTTATCAAAAGCTCTCAGGAAATCCGATATTC 60

Db 320 ATGAGACCCACCCGAGTTTGACCTTATCAAAAGCTCTCAGGAAATCCGATATTC 379
 QY 61 CTCTGCTTTTCCCAACAGAGTCAGGACATTCGGAGCTCCAGAGTGAAACCGGAGATG 120
 Db 380 CTCTGCTTTTCCCAACAGAGTCAGGACATTCGGAGCTCCAGAGTGAAACCGGAGATG 439
 QY 121 CTTGAGAGTTGGTGAATCGGAGTCGACAGAGATGCTGCGAGAGGCTCAGCGAG 180
 Db 440 CTTGAGAGTTGGTGAATCGGAGTCGACAGAGATGCTGCGAGAGGCTCAGCGAG 499
 QY 181 ACGGCAAGTAGGAGTATCGAAGTGCCTTGGAGTACATCAGTAAAGTGGCTACCTGGAC 240
 Db 500 ACGGCAAGTAGGAGTATCGAAGTGCCTTGGAGTACATCAGTAAAGTGGCTACCTGGAC 559
 QY 241 CCCAGGAATGAGAGATTTGCGAGTGCATCAAGAGACCTCCCGAGGAAGGCGCTGGCG 300
 Db 560 CCCAGGAATGAGAGATTTGCGAGTGCATCAAGAGACCTCCCGAGGAAGGCGCTGGCG 619
 QY 301 TCCACCCGCTGACTCGGCGCCAGTTTCGAGGGACAGGGGAAGCACTCCCATCTAC 360
 Db 620 CCCACCCGCTGACTCGGCGCCAGTTTCGAGGGACAGGGGAAGCACTCCCATCTAC 679
 QY 361 CACGAGCTGGTGGTGCAGAACTACGAGGGCCCGCCGCACTGGAGGAGATCCCGCGCAA 420
 Db 680 CACGAGCTGGTGGTGCAGAACTACGAGGGCCCGCCGCACTGGAGGAGATCCCGCGCAA 739
 QY 421 TATTTAGACTTTCTTCCCTGGAGCGGAGCTTTCGAGGGACAGGGGAAGCACTCCCATCTAC 480
 Db 740 TATTTAGACTTTCTTCCCTGGAGCGGAGCTTTCGAGGGACAGGGGAAGCACTCCCATCTAC 799
 QY 481 CATCTCCCAAGGTTACAGCAGCAGTACAGCAAGTGCAGCTTTCGCGGGCACACAC 540
 Db 800 CATCTCCCAAGGTTACAGCAGCAGTACAGCAAGTGCAGCTTTCGCGGGCACACAC 859
 QY 541 TATGGTGTGTCTACTACTATCGAGAGCTTGGGTATGGGGTGAGCGGAGTTCCTCC 600
 Db 860 TATGGTGTGTCTACTACTATCGAGAGCTTGGGTATGGGGTGAGCGGAGTTCCTCC 919
 QY 601 TTCAGAACAGAGCCGACAGATGCTTATTCAGAGTGGCAAGGCGGAGGCTGGCGCT 660
 Db 920 TTCAGAACAGAGCCGACAGATGCTTATTCAGAGTGGCAAGGCGGAGGCTGGCGCT 979
 QY 661 CCGCGAGCTCAGCTTTCTTCCCTGCGGAGTGTGGGTATGGGGTGAGCGGAGTTCCTCC 720
 Db 980 CCGCGAGCTCAGCTTTCTTCCCTGCGGAGTGTGGGTATGGGGTGAGCGGAGTTCCTCC 1039
 QY 721 GCGGCTACCCACCTGGGCGCCACCCATTACATGTTGGGACCCGCGGCTCCACGTTT 780
 Db 1040 GCGGCTACCCACCTGGGCGCCACCCATTACATGTTGGGACCCGCGGCTCCACGTTT 1099
 QY 781 ACTGGCGAAAGCTTCGACAGGCTGTGTCGACCGCTCCAGGAACAGGCTCAATGCTGAC 840
 Db 1100 ACTGGCGAAAGCTTCGACAGGCTGTGTCGACCGCTCCAGGAACAGGCTCAATGCTGAC 1159
 QY 841 TTGTACGAGCTGGGCTCCAGCGTCCGCTGGTGTGAGCTTCCAGTCCAGCGGCGGACTCG 900
 Db 1160 TTGTACGAGCTGGGCTCCAGCGTCCGCTGGTGTGAGCTTCCAGTCCAGCGGCGGACTCG 1219
 QY 901 CTGAGAGAGGAGGCTAGAGGCTTCGCGGCGGAGTGGCTTTTCGGGCTGGCGCCAGC 960
 Db 1220 CTGAGAGAGGAGGCTAGAGGCTTCGCGGCGGAGTGGCTTTTCGGGCTGGCGCCAGC 1279
 QY 961 AGGACCAACTCTTCAACAACCCCAACCTGAGCCCTCACTGCCCGCCCAACAGGTC 1020
 Db 1280 AGGACCAACTCTTCAACAACCCCAACCTGAGCCCTCACTGCCCGCCCAACAGGTC 1339
 QY 1021 ACCGCGGTGAGCGGCGACATCTTCACTGTGAGAGGCTGCTGTGCTGGGCGCC 1080
 Db 1340 ACCGCGGTGAGCGGCGACATCTTCACTGTGAGAGGCTGCTGTGCTGGGCGCC 1399
 QY 1081 GAGCCCCAGACAGCGTGGGCGCTTCGACCCGCTGGGCTGGGCTGGGCGCCACAGCACT 1140

Db 1400 GAGCCCCAGACAGCCGCTGGGGCCCTCGCACCCCGCTGGGTGGCTGCCCCACAGCACT 1459
 QY 1141 GGCAGTGAAGCTTGAGAGCAAGAGGAGGCGAGCGAGGCCACACCCGCTGGATGGGAC 1200
 Db 1460 GGCAGTGAAGCTTGAGAGCAAGAGGAGGCGAGCGAGGCCACACCCGCTGGATGGGAC 1519
 QY 1201 TATGGCGGCTCCGAGGCGAGGTCGCCACCGCTTCCCTATCCAAAGACATTCGCTGGCCC 1260
 Db 1520 TATGGCGGCTCCGAGGCGAGGTCGCCACCGCTTCCCTATCCAAAGACATTCGCTGGCCC 1579
 QY 1261 AGTAAGTCTGACAGTACAGCGTGGACCTGGACAGCTGTGCACCAAGTGTGCAGAGAGT 1320
 Db 1580 AGTAAGTCTGACAGTACAGCGTGGACCTGGACAGCTGTGCACCAAGTGTGCAGAGAGT 1639
 QY 1321 CTGCGAGGCGGCACTCTATAGACGGGAGTGTACAAAGAGCCAAAGGTGCGAAGGAGAG 1380
 Db 1640 CTGCGAGGCGGCACTCTAGCAAGAGAGAGTGTACAAAGAGCCAAAGGTGCGAAGGAGAG 1699
 QY 1381 AAAGCTGGGAGAGACAAAAGAGATTCAGACCTCCCGGTGCTTCCGCAAGATAGC 1440
 Db 1700 AAAGCTGGGAGAGACAAAAGAGATTCAGACCTCCCGGTGCTTCCGCAAGATAGC 1759
 QY 1441 AGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTCCTCCCTTATGCTTCAAAATTC 1500
 Db 1760 AGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTCCTCCCTTATGCTTCAAAATTC 1819
 QY 1501 TTGATGGAGCAACACGCTGGAGAAATGTATCAAAACCTTACAGCAGAGGTACGCGGAGG 1560
 Db 1820 TTGATGGAGCAACACGCTGGAGAAATGTATCAAAACCTTACAGCAGAGGTACGCGGAGG 1879
 QY 1561 CTACAGTGGAGGAGAAATGGCCAAAGCTGGGCTCTGTGAGGCGCGAGGAGGAGATG 1620
 Db 1880 CTACAGTGGAGGAGAAATGGCCAAAGCTGGGCTCTGTGAGGCGCGAGGAGGAGATG 1939
 QY 1621 AGAAGATCTCTTACCAGAAAGAGTCTTAACTACAAACCGCTGAAGAGGCGCAAGATGGAC 1680
 Db 1940 AGAAGATCTCTTACCAGAAAGAGTCTTAACTACAAACCGCTGAAGAGGCGCAAGATGGAC 1999
 QY 1681 AAGTCCATGTTTGTGAAAATCAAGACTCTAGGCAATGGGTCCTTTGGGGAAGTGTGCTC 1740
 Db 2000 AAGTCCATGTTTGTGAAAATCAAGACTCTAGGCAATGGGTCCTTTGGGGAAGTGTGCTC 2059
 QY 1741 GCTTGTAAAGTGGACACTCAGCTCTGTAGCCATGAAGACTCTCAGGAAGAGGATGTC 1800
 Db 2060 GCTTGTAAAGTGGACACTCAGCTCTGTAGCCATGAAGACTCTCAGGAAGAGGATGTC 2119
 QY 1801 CTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGAGATCTTGGCTGAAGCAGAC 1860
 Db 2120 CTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGAGATCTTGGCTGAAGCAGAC 2179
 QY 1861 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGAGGAGGCTTACTTTGTG 1920
 Db 2180 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGAGGAGGCTTACTTTGTG 2239
 QY 1921 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGAGGAGGCTTACTTTGTG 1980
 Db 2240 AATGAGTGGGTGTCAAACTCTACTCTTCCAGCAGAGGAGGAGGCTTACTTTGTG 2299
 QY 1981 GAGCAGCTGGGCGCTTCTACATTTGAGAGTGTGACCTTGGCCATTTGAAGTGTCCCAAG 2040
 Db 2300 GAGCAGCTGGGCGCTTCTACATTTGAGAGTGTGACCTTGGCCATTTGAAGTGTCCCAAG 2359
 QY 2041 ATGGGCTTTATCCACCGGAGATCAAGCTTGAACATCTATCTGAGCTTGGATGGTGTGAT 2100
 Db 2360 ATGGGCTTTATCCACCGGAGATCAAGCTTGAACATCTATCTGAGCTTGGATGGTGTGAT 2419
 QY 2101 ATTAAGCTGACAGATTTTGGCTCTGCACTGGATTCAGGTGGATTCACAAATTCAGATAC 2160
 Db 2420 ATTAAGCTGACAGATTTTGGCTCTGCACTGGATTCAGGTGGATTCACAAATTCAGATAC 2479
 QY 2161 TACAGAAAGGAGACCATGAGCAGGAGCAGATGAGCGCGGCTGACCTCTGGGAGGAT 2220
 Db 2480 TACAGAAAGGAGACCATGAGCAGGAGCAGATGAGCGCGGCTGACCTCTGGGAGGAT 2539


```

QY 2221 GTTTCACCTGCGCTGTGGAGACAGTTAAAGACCTTGAGCAGAGGCGCGAAGACG 2280
Db 2540 GTTTCACCTGCGCTGTGGAGACAGTTAAAGACCTTGAGCAGAGGCGCGAAGACG 2599
QY 2281 CACCAGAGTGCCTGCACATCTCTGTGGGACACCAAAATACATCGTCGGAGGTG 2340
Db 2600 CACCAGAGTGCCTGCACATCTCTGTGGGACACCAAAATACATCGTCGGAGGTG 2659
QY 2341 CTTCTCCGCAAGGTTACAGCAGCTCTGTACTGTGGAGCTCGGTGATTCTCTTT 2400
Db 2660 CTTCTCCGCAAGGTTACAGCAGCTCTGTACTGTGGAGCTCGGTGATTCTCTTT 2719
QY 2401 GAGATCCTGTTGGGAGCGCGCTTTCTTGGCCCCCACCACACAGAGCAGCTGAAG 2460
Db 2720 GAGATCCTGTTGGGAGCGCGCTTTCTTGGCCCCCACCACACAGAGCAGCTGAAG 2779
QY 2461 GTGATCACTGGGAGAGCAGCTGTGCATATCCCTACGAGGTGAGGCTGAGGCG 2520
Db 2780 GTGATCACTGGGAGAGCAGCTGTGCATATCCCTACGAGGTGAGGCTGAGGCG 2839
QY 2521 CGAGACCTCATACAGAGCTGTGCTCGCGGCTGACTGCGCTGGCAGGATGGGCA 2580
Db 2840 CGAGACCTCATACAGAGCTGTGCTCGCGGCTGACTGCGCTGGCAGGATGGGCA 2899
QY 2581 GATGACCTCAAGGCACACCGCTTCTTCAACACCATCGACTTTTCCGCTGACATCCGAAAG 2640
Db 2900 GATGACCTCAAGGCACACCGCTTCTTCAACACCATCGACTTTTCCGCTGACATCCGAAAG 2959
QY 2641 CAGGCTGACACCTTACCTCCCGACCAATCAGCAGCCCATGACACCTTCCAAATTTGACCG 2700
Db 2960 CAGGCTGACACCTTACCTCCCGACCAATCAGCAGCCCATGACACCTTCCAAATTTGACCG 3019
QY 2701 GTGGATGAGAAAGCCCTTGGACAGAGCGCCGAGAGAGCGCCAAAGCCTGGGACAG 2760
Db 3020 GTGGATGAGAAAGCCCTTGGACAGAGCGCCGAGAGAGCGCCAAAGCCTGGGACAG 3079
QY 2761 CTGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTCACTTCCGACAG 2820
Db 3080 CTGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTCACTTCCGACAG 3139
QY 2821 TTCTTCGATGACAAAGCTTATCCCTTCCGCTGCGGAGCGCTTCAAGCCCTCAGAGCCCGCAGAGT 2880
Db 3140 TTCTTCGATGACAAAGCTTATCCCTTCCGCTGCGGAGCGCTTCAAGCCCTCAGAGCCCGCAGAGT 3199
QY 2881 GCAGACCCAGGGATCGGACTTGGAGCTGGGCGGAGGCTGCGGCGGCTGACGCTGACGTTG 2940
Db 3200 GCAGACCCAGGGATCGGACTTGGAGCTGGGCGGAGGCTGCGGCGGCTGACGCTGACGTTG 3259
QY 2941 TAAGCCTCAGTTAAACCAACTCGAGGAAACCAAAATGAGATTTCTTTTCAGAAACAA 3000
Db 3260 TAAGCCTCAGTTAAACCAACTCGAGGAAACCAAAATGAGATTTCTTTTCAGAAACAA 3319
QY 3001 ACTCAAGCTTAGAATCCCTTCAATTTTGTGTTAGTCTGTTAAATGGCAACAGAGAGTCAAC 3060
Db 3320 ACTCAAGCTTAGAATCCCTTCAATTTTGTGTTAGTCTGTTAAATGGCAACAGAGAGTCAAC 3379
QY 3061 ATGATTTCAATTAGCCCTCTGAGGACCTTCACTGATTAACACAGTATTTTAAAAA 3120
Db 3380 ATGATTTCAATTAGCCCTCTGAGGACCTTCACTGATTAACACAGTATTTTAAAAA 3439
QY 3121 TTAGTACAGTATGGAAGA 3139
Db 3440 TAAAAA

```

```

RESULT 3
E38226
LOCUS E38226 5486 bp DNA linear PAT 31-JAN-2002
DEFINITION Human tumor regulatory gene.
ACCESSION E38226
VERSION E38226.1 GI:18626935
KEYWORDS JP 2000210086-A/1.

```

```

SOURCE Homo sapiens.
ORGANISM Homo sapiens
REFERENCE Koga, J., Kono, K. and N.Z.F.
AUTHORS Human tumor regulatory gene
TITLE Patent: JP 2000210086-A 1 02-AUG-2000;
JOURNAL NIHON CHEMICAL RESEARCH K K
COMMENT OS Homo sapiens (human)
PN JP 2000210086-A/1
PD 02-AUG-2000
PF 25-JAN-1999 JP 1999016223
PR
PI JUNICHI KOGA, KEIKO KONO, ZOROTARYOFU FIODORU N PC
C12N15/09,A61K31/00,A61K38/27,C07K14/47,C07K14/525,C12Q1/68, PC
G01N33/50,
PC G01N33/50,G01N33/574,G01N33/577//C12P21/08,(C12N15/09, PC
C12R1:91),C12N15/00,
PC A61K37/36,(C12N15/00,C12R1:91)
CC
FH Key Location/Qualifiers
FT source 1. 5486
/organism='Homo sapiens'
/organism='Homo sapiens'
/db_xref='taxon:9606'
BASE COUNT 1378 a 1452 c 1390 g 1266 t
ORIGIN
Query Match 55.5%; Score 1749.6; DB 6; Length 5486;
Best Local Similarity 77.4%; Pred. No. 0;
Matches 2389; Conservative 0; Mismatches 544; Indels 153; Gaps 16;
QY 1 ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAAGCTCTCAGGAAATCCGATATTC 60
Db 576 ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAAGCTCTCAGGAAATCCGATATTC 635
QY 61 CTCCTGCTTTTGGCCAAACAGTCAGGCACCTTCGGAGCTGCGAGGTGAACCGGAGATG 120
Db 636 TTGTTGCTTTTGGCTTAATGAATCGGACCTCTGAGCTGCGAGAGTGAACCGGAAATG 695
QY 121 CTTGAGAGCTTGTGAATCGGCATGTGACAGGAGATGGCTGGCAGAGCGTCAAGCAG 180
Db 696 CTTGAGAGCTTGTGAATCGGCATGTGACAGGAGATGGCTGGCAGAGCTCAAGCAG 755
QY 181 ACGGCAGTAGGATGATCGAAGCTGCTTTGGAGTACATCAGTAAGATGGCTTACCTGGAC 240
Db 756 ACTGGCAGCAGGAGCATCGAGCGGCTGGAGTACATCAGCAAGATGGCTTACCTGGAC 815
QY 241 CCGAGGATGAGCAGATTTGTGGAGTCAATCAAGCAGACCTCCCGAGGAAGGCGCTGGCG 300
Db 816 CCGAGGATGAGCAGATTTGTGGAGTCAATTAAGCAGACCTCCCGAGGAAGGCGCTCATG 875
QY 301 TCACCCCGGTCGCTCGGCGGCCAGTTTCGAGGGGACACAGGGAAGCACTCCCATCTAC 360
Db 876 CCAACCCGAGTACGCGGAGGCGCCAGCTTCGAGGAGNACCGGCGATTCGTTTGGTCTAC 935
QY 361 CACCACTGGGTGGTGGCAAACTACGAG-----GGCCCCCGCGCA 399
Db 936 CACCACTGAGCGGTACCCCTTACGAGGCGCCCAAGCTTCGCGCGTGAACGCGCCGAGGG 995
QY 400 CTGGAGGAGATGCGCGGCGCAATATTAGACTTTCTTCCCTGGAGCGCGGAGCGGAC 459
Db 996 CTGGAGGAGATGCGCGGCGCTACCTGGAGTACCTTTTCCCGGAGTTCGCGCCACG 1055
QY 460 CACGTTGCCAGGCTCACCAGCATCTCCCAAGGTTACAGCAGCAGCAGTAGAGCAAGT 519
Db 1056 CCGGCGCAGCAG---CACCAGCAGCCCAAGGCGTACGTTCCCGCAGCAGCAGCAGCA 1112
QY 520 GCG-----CACTTTCGGGCGCACACTATGTTGTTGTTGTTCTACT-----ATCG 564

```

Db 1113 GGGGACACATTCCTCCGCTGACAGGGCGCGCACTACGCGGCGCGCACCTGTGTGCTGGG 1172
QY 565 GAGCAGTCTGGGTATGGGTGACGCGAGTTCCTCTCTCCAGAACAAAGACGCGCACAGAT 624
Db 1173 GAACCCCTGGGTACGAGGTGACGCGACGCGCTCTCTCCAGAGCAAGACGCGCGGAG 1232
QY 625 GCC-----TATTCACAGATGGC---AAGGCCAGGGTGGCCCTCC---CGCCAGCCCTC 672
Db 1233 ACCGGGGGTATGCCAGCTGCCACGAAAGGCGAGGAGGACCGCGAGCGCGCGCTC 1292
QY 673 ACCTTTCTCTGCCCATCTGGGTGTACACTGCTGCTGCACCAACAAGCGCGGCTACCCCA 732
Db 1293 GCTTTCCACCCCTGCCCGGGCTCTAGTGCACCAACCAAGCAGCGCGGT 1352
QY 733 CTGGGGCCACCAATACATGTGTGGCACCGGGTCCACAGTTTACTGGCGAAAGC 792
Db 1353 CCGCGGGCCACAGCTGATGTGTGGCTCCCGCAGCAGGTGTTCGCCAGCGACAGC 1412
QY 793 TCTGACAGGCTGTGCTGCGACCGGTCCAGGAACAGCTCAATGCTGACTGTACAGCTG 852
Db 1413 CCCCAGAGGCTGTCTACTCCCTCGGGGAACAGCTCAAGCTGGAGCTGTATGAATTG 1472
QY 853 GG-----CTCCACAGGCTGGCTGTGCTGAGCTCCACTGGCACGCGCGGACTTCGCTGCG 906
Db 1473 GCGAGACCTCTCGTCCAGAGTGGCGGCTGCCACCTTGGCGCGCGGACTCCCTGCGAG 1532
QY 907 AAGCAGGCTTGAAGACCTTCGGGCG---CGCATGTGGTTCGCGGCTGGC----- 954
Db 1533 AAGCGGGGCTGGAGGCGCGCGCGGCGACGTGGCTTCGGGCTGACTGCCAGTG 1592
QY 955 CCACAGAGCACTCTCTCAACAAACCAAC-----CT 990
Db 1593 CCGCAGAGCACTCTCAACAGCCACAGCGCGCGCGCTGCCCTGGCAAGGCC 1652
QY 991 GAGCCCTCACTGCCCGCCCAACAGCTCACCGCGGTGACGGCGGACACATCTTCCAC 1050
Db 1653 GAGCCCTCCCTGCCCGCCCAACAGCTGACGGCTGTACGGCGCGGACATCTTGAC 1712
QY 1051 CCGTGAAGAGGCTGCTGCTGGCGCGGAGCGGCCAGCGCGGTGGGGCCCTCGCAC 1110
Db 1713 CCGTGAAGAGGCTGCTGCTGAGGCGCGGAGCGCGGTGGGGCCCTCGCAC 1772
QY 1111 CCGCCTGGGTGGCTGCGC-----CCACAGCACTGCGCACATGAG 1149
Db 1773 CCGCCTGGGTGGCGGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCTGCGGAG 1832
QY 1150 AGCCTGGAGACGAGGA-----GGGACGCGAGCGCCACACCCGCTG 1191
Db 1833 GCGTGGAGCGCAAGGAGGAGCATGCGCTGGCGTGGCGGCGGCGCGCTTCCGCGTG 1892
QY 1192 GATGTGACTATGGGCTCCGAGCGGAGTGGCCCGCGCTCGTATCCAAGCACTTG 1251
Db 1893 GACGTGAGTACGAGGCGCGGAGCGGAGTGGCGCGCTCCGCGCTACCGGAGACCTG 1952
QY 1252 CTGCTGCCAGTAAGTCTGAGCAGTACAGGCTGGAACCTGGACAGCCTGTGCACAGTGTG 1311
Db 1953 CTGCTGGCAGCAAGTCGAGCAGTAC-----GACCTGACAGCCTGTGCGGAGGATG 2006
QY 1312 CAGCAGAGTCTGCGAGGGG---CACTGTATCTAGACGGGAGTGACAGGCCACAAAGGT 1368
Db 2007 GAGCAGAGCTCTGCTGGGCGCGGAGCGGAGTGGCGCGCTCCGCGCTACCGGAGACCTG 2066
QY 1369 GCGAAGGAGCAAGCTGGCAGAGCAAAAGCAGATTCAGACCTCCCGGCTGCTGTC 1428
Db 2067 GCGAAGGAGCAAAAGCGGAAAGGATAAAGAGCATTCAGACCTCTCCCGTTCGCGT 2126
QY 1429 CGCAAGATAGCAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTACTCCGCTTAT 1488
Db 2127 CGCAAAACAGCAGAGAGAGAGAGAGTACAGCATCAAGAGTACTCGGCATAC 2186
QY 1489 GCCTTCAATCTTCATGGAGCAACAGGTGGAGAAATGTATCAAAACCTTACAGCAGAAAG 1548
Db 2187 GCCTTAAAGTCTTCATGGAGGAGCAGCTGGGAATGTATCAAAACCTTACAGCAGAAAG 2246

QY 1549 GTACGCGGAGGCTACAGCTGGAGCAGAAATGGCCAAAGCTGGCTCTGTGAGGCCGAG 1608
Db 2247 GTTAACCGGAGGCTCAGCTGGAGCAAGAAATGGCCAAAGCTGACTCTGTGAAGCTGAG 2306
QY 1609 CAGGAGCAGATGAGGAAGATCCTCTACCAAGAGAGTCTAACTACAACCGCTGAAGAGG 1668
Db 2307 CAGGAGCAGATGCGGAAGATCCTCTACCAAGAGAGTCTAATTAACAGGTTAAAGAGG 2366
QY 1669 GCGAAGATGGACAAAGTCCATGTTGTGAAATCAAGACTCTTAGCATCGGTGCTTTGGG 1728
Db 2367 GCGAAGATGGACAAGTCTATGTTGTCAAGATCAAAACCTTGGGATCGGTGCTTTGGA 2426
QY 1729 GAAGTGTCCCTCGCTGTAAAGCTGAGACTCAGCTCTGTAGCCATGAAGACTCTCAGG 1788
Db 2427 GAAGTGTCCCTCGCTGTAAAGCTGAGACTCAGCTCAGCCCTGTACGCCATGAAGACTTAAG 2486
QY 1789 AAGAGGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGAGCATCTG 1848
Db 2487 AAAAGGATGCTCTGAACCGGAATCAGTGGCCCATCAGTGGCCAGCGAGAGGGAGCATCTG 2546
QY 1849 GCTAAGCAGACAATGAGTGGGTGCTCAAACTCTACTCTCTTCCAGGACAAAGCAGC 1908
Db 2547 GCCAGGACAGACAATGAGTGGGTGCTCAAACTCTACTCTCTTCCAGGACAAAGCAGC 2606
QY 1909 CTGTACTTGTGATGGACTACATACAGGCGGGGATATGATGAGCCTGCTGATCAGGATG 1968
Db 2607 CTGTACTTGTGATGGACTACATCCCTGGTGGGACATGATGAGCCTGCTGATCCGAGT 2666
QY 1969 GAGTCTTCCCTGAGCACCTGGCCGCTTCTACATTCAGAGATTGACCTGGCCATTTGAA 2028
Db 2667 GAGTCTTCCCTGAGCACCTGGCCGCTTCTACATCCAGAGCTGACTTTGGCCATTTGAG 2726
QY 2029 AGTGTCCACAAGTGGCTTTATCCACCGGAGCATCAAGCCTGACAACTACTCTATCGAC 2088
Db 2727 AGTGTCCACAAGTGGCTTTATCCACCGGAGCATCAAGCCTGACAACTACTCTATCGAC 2846
QY 2089 CTGATGGTCTATTAAGCTGACAGATTTTGGCTCTGCACTGGATTTCAGTGGGACTCAC 2148
Db 2787 CTGATGGTCTATTAAGCTGACAGATTTTGGCTCTGCACTGGATTTCAGTGGGACTCAC 2846
QY 2149 AATTCAGTACTACCAAGAGGACCATGAGCAGAGCAGCATGGAGCCCGGTGAC 2208
Db 2847 AATTCAGTACTACCAAGAGGACCATGAGCAGAGCAGCATGGAGCCCGGTGAC 2906
QY 2209 CTCTGGCAGCATGTTTCCAACTGCTGCTGTGGAGACAGGTTAAAGACCTGGAGAGAGG 2268
Db 2907 CTCTGGGATGATGTCTAACTGCTGCTGGTGGGACAGGCTGAAGACCTTAGAGCAGAGG 2966
QY 2269 GCGCAGAGCAGCAGCAGAGGTCCTGGCACATTTCTTGTGGGACACCAATTAATC 2328
Db 2967 GCGCGAAGCAGCAGCAGAGGTCCTGGCACATTTCACTGGTGGGACTCCAACTACATC 3026
QY 2329 GCTCCCGAGGTCCTTCTCGCAAGGTTACAGCAGCTCTGTGACTGTGGGAGCGTCTG 2388
Db 3027 GCACCCGAGTCTCTCTCCGCAAGGTTACTCAACTCTGTGACTGTGTGGAGTGTGGA 3086
QY 2389 GTGATTTCTTTTGTGAGATGCTGTTGGGACGCGCTTCTTGTGGGACACCAATTAATC 2448
Db 3087 GTGATTTCTTTTGTGAGATGCTGTTGGGACGCGCTTCTTGTGGGACACCAATTAATC 3146
QY 2449 AGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGCAATATCCCTACGAGGTGAGGCTC 2508
Db 3147 ACCCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGCAATATCCAGCCAGGTGAGCTG 3206
QY 2509 AGCGCTGAGGCGGAGACCTCATCAGAGCTGTGCTGGCGGCTGACTGCGCCTGGGC 2568
Db 3207 AGCGCTGAGGCGGAGACCTCATCAGAGCTGTGCTGGCGGCTGACTGCGCCTGGGC 3266
QY 2569 AGGATGGGCGGAGTGAACCTCAAGGACACACCGTTCTTCAACACCATCGACTTTTCCGT 2628
Db 3267 CGGAATGGGCGGAGTGAACCTCAAGGACACACCGTTCTTCAAGGCGGCTTCTTCAAGGCGG 3326

```

QY 2629 GATATCCGAAAGCAGGCTGCACCTACGCTCCACCATCAGCACCACCCATGGACACCTCC 2688
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3327 GATATCCGAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3386
QY 2689 AATTTTGACCGGTGATGAAGAAGCCCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2748
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3387 AATTTTGACCGGTGATGAAGAAGCCCTGGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 3446
QY 2749 GCCTGGGACACGCTGGCTCCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2808
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3447 GCCTGGGACACGCTGGCTCCGCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3506
QY 2809 ACCTTCCGAGGTTCTTCGATGACAAAGCCTATCCCTCCCGTGGCCGAGGAGGAGGAGGAG 2866
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3507 ACCTTCCGAGGTTCTTCGATGACAAAGCCTATCCCTCCCGTGGCCGAGGAGGAGGAGGAGG 3566
QY 2867 -----ACCCGCGAGAGTGCAGACCCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2916
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3567 GCAGAAGCTTCACAGGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGGTGGATCAGACT 3626
QY 2917 GAGGCTGCACGCGGTGTACGTGTA 2942
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
3627 GAAGCTGCCACGCTGTGTACGTGTA 3652
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

RESULT 4
E38227 5486 bp DNA linear PAT 31-JAN-2002
LOCUS Human tumor regulatory gene.
DEFINITION E38227
ACCESSION E38227
VERSION 1
KEYWORDS JP 200210086-A/2.
SOURCE Homo sapiens.
ORGANISM Homo sapiens.
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 5486)
AUTHORS Koga, J., Kono, K. and N. Z.F.
TITLE Human tumor regulatory gene
JOURNAL Patent: JP 200210086-A 2 02-AUG-2000;
NIHON CHEMICAL RESEARCH K K
COMMENT OS Homo sapiens (human)
PN JP 200210086-A/2
PD 02-AUG-2000
PF 25-JAN-1999 JP 1999016223
PR
PI JUNICHI KOGA, KEIKO KONO, ZOROTARYOFU FIODORU N PC
C12N15/09,A61K31/00,A61K38/27,C07K14/47,C07K14/525,C12Q1/68, PC
G01N33/50,
PC G01N33/50,G01N33/574,G01N33/577//C12P21/08,(C12N15/09, PC
C12R1:91),C12N15/00.
CC A61K37/36,(C12N15/00,C12R1:91)
FH Key Location/Qualifiers
FT source 1..5486
FT /organism="Homo sapiens (human)".
FEATURES
source 1..5486
BASE COUNT 1380 a 1451 c 1388 g 1267 t
ORIGIN
Query Match 55.4%; Score 1748; DB 6; Length 5486;
Best Local Similarity 77.4%; Pred. No. 0;
Matches 2388; Conservative 0; Mismatches 545; Indels 153; Gaps 16;
QY 1 ATGAGGACACCCCGGAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCCGATATTC 60
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
576 ATGAGGACACCCCGGAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCCGATATTC 635
QY 61 CTCCTGCCTTTTGGCAACAGTTCAGGCACTTCGGCAGCTGCAGAGTGAACCGGAGATG 120
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||

```

```

Db 636 TTGTTGCCCTTTTGTAAATGAATCGGCGACCTCTGCAGCTGCAGAGTGAACCGGCAATG 695
QY 121 CTTTCAGGAGTTTGGTGAATCGGCGCATGTGACACGAGGAGATGGCTGGCAGAGCGCTCAGCGAG 180
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
696 CTGCAAGAACTTGGTGAACGCGAGGATGCGACGACGAGGATGGCTGGCGGAGCTCTCAAGCAG 755
QY 181 ACGGGCAGTAGGAGTATCGAAGCTGCTTGGAGTACATCAGTAAGATGGCTTACCTGGAC 240
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
756 ACTGGCAGCAGGAGATCGAGGCGCCCTGGAGTACATCAGCAAGATGGCTTACCTGGAC 815
QY 241 CCAGAGATGACAGATGTGAGTATCAAGCAGACCTCCCGAGGAAGGCGCTGGCG 300
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
816 CCGAGGATGACAGATTTGCGGGTCAATTAAGCAGACCTCCCGAGGAAGGCGCTCATG 875
QY 301 TCCACCCCGTCACTCGGCGCCAGTTTCGAGGCGACAGGAGGAGCACTCCCATCTCTAC 360
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
876 CCAACCCCGTCACTCGGCGGAGGCCAGCTTCGAAGGAACCGCGGATTCGTTTCGCTCTAC 935
QY 361 CACAGCTGGTGGTCAAACTACGAG-----GGCCCGCCGCA 399
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
936 CACAGCTGAGCGGTACCCCTACGAGGCGCCCAAGCTTCGGCGCTCAGCGCCCGCAGGCG 995
QY 400 CTGGAGGAGATGCGCGGCAATATTAGACTTTCTTCCCTGGAGCGGAGCGCGGACAC 459
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
996 CTGGAGGAGATGCGCGGCGCTACGTGGACTACCTTTCCCGGAGTCCGCCCGCCACGGG 1055
QY 460 CAGGTGCCAGGCTCACGAGCATCTCCCAAGGTTACAGCAGCAGCAGTACAGCAAGT 519
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1056 CCGCGCCACGAG---CACCAGCAGCCACCAAGGGCTAGGGTGCCAGGCTAGAGGCGAGCA 1112
QY 520 GCG-----CACTTTCCGGGCGACACACTATGTGCTGGTCACTACT-----ATCG 564
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1113 GGGGCGACACTCCCGCTGCGAGGCGCGCACTAGGCGGCGCGCACTGCTGGTCTGGG 1172
QY 565 GAGCAGTGTGGTATGGGTGCGAGCGAGTCTCTCTTCCAGAAAGACGACGACCATGAT 624
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1173 GAACCCCTGGGTACGGAGTGCAGGCGAGCCCTCTCTTCCAGAGGAAGACGCGCGGAG 1232
QY 625 GCC-----TATTCAGAGTGGC---AAGGCCAGGGTGGGCCCTCC---CGCAGCCTC 672
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1233 ACCGGGGTTACGCGAGCTGCCACGAAGGGCCAGGAGGACCGCGAGCGCGCGCTC 1292
QY 673 ACCTTTCTGCCCCATGCTGGGCTGTACACTGCTCGCACCAAGCGCGGCTACCCCA 732
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1293 GCTTTCCCGACCCCTGCGCGCGGCTCTACGTGCGCGCACCAACACAGCAGCGCGT 1352
QY 733 CTGGGGCCCGCCCATATGATGTGTGGCAGCCCGGGTTCACAGTTTACTGGCGAAAGC 792
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1353 CCGGTGGCGCCACAGCTGCTGTGGCTCCCGCAGCAGGTGTTCGCCAGCGACAGCAGC 1412
QY 793 TCTGCACAGGCTGTGCTGCGACCGTCCAGGAACAGCCTCAATGCTGACTGTACGAGCTG 852
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1413 CCGCGCAGAGCTGCTCACTCCCTCGGGAACAGCCTCAACGCTGACTGTATGAATG 1472
QY 853 GG-----CTCCAGGCTGCGCTGGTGTGAGCTCCACTGGCACGCGGCGACTTCGCTGAG 906
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1473 AGCAGCAGCTCGTCCAGCAGTGGCGGCTGCCACCTGGCGCGCGGCGACTTCCTGAG 1532
QY 907 AAGCAGGCTCTAGAAGCTTCGCGG---CGCATGTGGCTTTTCGGGCTGGC-----954
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1533 AAGCGCGCTTGGAGGCGCGCGCGCGCGCGCTTCCGGCCTGACTGCGCGCAGTG 1592
QY 955 CCAGCAGGACCACTCTCTCAACACCAACACAC-----CT 990
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1593 CCAGCAGGACCACTCTCTCAACAGCAGCAGCGCGCGCGCGCTCGCGCTGGCAGGCGC 1652
QY 991 GAGCCTCACTGCCCGCCCGCCCAACACAGGCTACCGCGCTGACGCGCGCACATCTTCAC 1050
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1653 GAGCCTCCCTGCCCGCCCGCCCAACACAGGCTGACGCGCTGTACGCGCGCGCACATCTTCAC 1712
QY 1051 CTTGTGAGAGGCTGCTGTGCTGGCGCGGAGCGCCAGACAGCGGTGGGCGCTTCGAC 1110
Db      ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
1713 CCGGTGAAGAGGCTGCTGTGCTGAGGCGGAGCGCGCGCGCTTCCGGCCTGACTGCGCGC 1772

```

Qy 1111 CCCGCTGGGTGGTGGC-----CCACAGCACCTGCCACTGAG 1149
 Db 1773 CCCGCTGGGTGGTGGC-----CCACAGCACCTGCCACTGAG 1832
 Qy 1150 AGCTGGAGACGAAGG-----GGCAGCGGAGGCCACACCCGCTG 1191
 Db 1833 GCTTGGACGCAAGGAGGAGCATGCCCTGGCTGGCGGCGAGGCCCTTCCCGCTG 1892
 Qy 1192 GATGTGGACTATGGCGCTCCGAGCGCAGGTGCCACCGCTCCGTATCCAAAGCACTTG 1251
 Db 1893 GACGTGGAGTACGAGGCCACAGACCGGAGGTGCCCGCTCCGCCCTACCCGAAAGCACCTG 1952
 Qy 1252 CTGCTGCCAGTAAAGTCTGAGCAGTACAGCGTGGACCTGGACAGCCTGTGCACCACTGTG 1311
 Db 1953 CTGCTGCCAGCAAGTCGAGCAGTAC-----GACCTGGACAGCCTGTGCGCAGGCATG 2006
 Qy 1312 CAGCAGAGTCTGCGAGGGG---CACTGATCTAGACGGAGTACACAGAGCCACAAAGGT 1368
 Db 2007 GAGCAGAGCTTCCGTGGCGGCCCAACAGAGCCCGAGGCGCGACAGAGCGCAAAAGC 2066
 Qy 1369 GCGAAGGAGACAAGCTGGCAGAGACAAAAGCAGATTACAGACTCCCGCGTGCCTGTC 1428
 Db 2067 GCCAAGGGGACAAAGCGGAAAGATATAAAGCAGATTACAGACTCTCCCGTTCGCGTC 2126
 Qy 1429 CGCAAGATAGCAGAGATGAAGAGAAGAGAGTCTCGCATCAAGAGTACTCCCTTAT 1488
 Db 2127 CGCAAAAACACAGAGAGAGAGAGAGAGTCAAGCATCAAGAGTACTCGCCATAC 2186
 Qy 1489 GCCTTCAATTTCTATGAGCAACACAGTGGAGATGTATCAAAACCTACACAGAG 1548
 Db 2187 GCCTTAAAGTTCTTATGAGCAGCAGTGGAGATGTATCAAAACCTACACAGAG 2246
 Qy 1549 GTCAGCGGAGGTACAGCTGAGCAGAGAAATGCCAAAGCTGGCTGTGAGGCGCAG 1608
 Db 2247 GTTAAACGGAGCTGACAGTGGAGCAAGAAATGCCAAAGCTGGACTGTGAAAGCTGAG 2306
 Qy 1609 CAGGAGCAGATGAGGAAGATCCTTACCAGAAAGAGTCTTAACAAACCGCTGAAGAGG 1668
 Db 2307 CAGGAGCAGATGCGGAAGATCCTTACCAGAAAGAGTCTTAACAAACCGCTGAAGAGG 2366
 Qy 1669 GCCAAGATGACAGATCCATGTTGTGAAATCAAGACTCTAGCATCGGTGCTTTGGG 1728
 Db 2367 GCCAAGATGACAGATCCATGTTGTGAAATCAAGACTCTAGCATCGGTGCTTTGGG 2426
 Qy 1729 GAAGTGTCCCTGCTGTGAAGCTGACACTCACGCTCTGTACGCCATGAAGACTCTCAGG 1788
 Db 2427 GAAGTGTCCCTGCTGTGAAGCTGACACTCACGCTCTGTACGCCATGAAGACTCTCAGG 2486
 Qy 1789 AAGAAGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 1848
 Db 2487 AAGAAGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 2546
 Qy 1849 GCTGAACAGACATGAGTGGTGGTCAAACTCTACTCTCTTCCAGGACCAAGGACAGC 1908
 Db 2547 GCCAGGACAGACATGAGTGGTGGTCAAACTCTACTCTCTTCCAGGACCAAGGACAGC 2606
 Qy 1909 CTGTACTTTGTGATGAGTACATACAGGCGGGATATGATGAGCCTGCTGATCAGGATG 1968
 Db 2607 CTGTACTTTGTGATGAGTACATACAGGCGGGATATGATGAGCCTGCTGATCAGGATG 2666
 Qy 1969 GAGTCTTCCCTGAGCACCTGGCCCGTCTTACATTTGACAGTTGACCCCTGGCCATTGAA 2028
 Db 2667 GAGTCTTCCCTGAGCACCTGGCCCGTCTTACATTTGACAGTTGACCCCTGGCCATTGAG 2726
 Qy 2029 AGTCTCACAGATGGGCTTTATCCACCGGACATCAAGCCTGACACATACTCATGAC 2088
 Db 2727 AGTCTCACAGATGGGCTTTATCCACCGGACATCAAGCCTGACACATTTGATAGAT 2786
 Qy 2089 CTGATGGTCTATTAAGCTGACAGATTTTGGCTCTGCACTGCTGATTCAGTGAGCTCAC 2148
 Db 2787 CTGATGGTCTATTAAGCTGACAGATTTTGGCTCTGCACTGCTGATTCAGTGAGCTCAC 2846

Qy 2149 AATTCCAAGTACTACCAAGAAAGGAAACCATGAGACAGGACATGAGACAGGACCGCGTGAC 2208
 Db 2847 AATTCCAATATATACCAAGAAAGGAGCCATGTACAGACAGGACAGCATGAGACCGCAGAC 2906
 Qy 2209 CTCTGGGACGATGTTTCCAACTGTCGTGGACACAGGTTAAAGACCCCTGGAGCAGAGG 2268
 Db 2907 CTCTGGGATGATGTCTAACTGTCGTGGGACAGGCTGAAGACCCCTAGACAGAGG 2966
 Qy 2269 GCGCAGAAAGCAGACACAGAGGTGCTGSCACATCTCTTGTGCGGACACCAAAATACATC 2328
 Db 2967 GCGCGAAGCAGACACAGAGGTGCTGSCACATCTCACTTGTGGGACCTCCAAACTACATC 3026
 Qy 2329 GCTCCGAGGTGCTTCTCCGCAAAAGGTACACGAGCTCTGTGACTGTGTGAGGAGGTCGGT 2388
 Db 3027 GCACCCGAGGTGCTTCTCCGCAAAAGGTACACTCAACTCTGTGACTGTGTGAGTGTGGA 3086
 Qy 2389 GTGATTTCTTTGAGATGCTGTTGGGACGCGCTTTCTTGGCCCCCACCACACAGAG 2448
 Db 3087 GTGATTTCTTTGAGATGCTGTTGGGACGCGCTTTCTTGGCACCTACTCTCCACAGAA 3146
 Qy 2449 ACGCAGCTGAAGGTGATCACTGGGAGACGCTGCATATCCCTACGAGTGGAGCTC 2508
 Db 3147 ACCCAGCTGAAGGTGATCACTGGGAGACGCTGCATATCCAGCCAGGTTGAAGCTG 3206
 Qy 2509 AGCCTGAGGCGCCGAGACCTCATCAGAAAGCTGTGCTGCGGGGTGACTGCCGCTGGGC 2568
 Db 3207 AGCCTGAGGCGCCGAGACCTCATCAGAAAGCTGTGCTGCGGAGACCCAGCTGGG 3266
 Qy 2569 AGGATGGGCGAGATGACCTCAAGGCACACCCGTTTCTTCAACACCATCGACTTTCCGCT 2628
 Db 3267 CGAATGGGCGGATGACCTGAAGGCCACCCCTTCTTCAAGCCATTCAGCTTCTCCAGT 3326
 Qy 2629 GACATCCGAAAGCAGGCTGCACCTACGTCGCCACCATCAGCCACCCATCGAGCACCTCC 2688
 Db 3327 GACATCCGAAAGCAGGACGACCCCTACGTCGCCACCATCAGCCACCCATCGAGCACCTCG 3386
 Qy 2689 AATTTGACCCGCTGGATGAAGAAAGCCCTTGGCACAGGCGCAGCGAGAGAGCGCAAG 2748
 Db 3387 AATTTGACCCGCTGGATGAAGAAAGCCCTTGGAAACGATGCCAGGAAGTAGCACCAAG 3446
 Qy 2749 GCCTGGACACGCTGGCTCCCGCAGCAGCAAGCATCCAGAGACGCGCTTCTATGAGTTC 2808
 Db 3447 GCCTGGACACACTCACCTCGCCCAATAAACAGCATCCTGAGCAGCATTTTACGAATTC 3506
 Qy 2809 ACCTTCGCGAGTCTTTCGATGACAGCGCTATCCCTTCGCTGCGGTCGCCAGCCCTCAG-- 2866
 Db 3507 ACCTTCGCGAGTCTTTCGATGACAGCGCTATCCCTTCGATGCCCAAGCCCTTCAGGA 3566
 Qy 2867 -----AGCCCGCAGAGAGTGCAGACCCAGGGGATCGGAGCTTGGAAAGTGGCGCC 2916
 Db 3567 GCAGAAGCTTCACAGGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGGTGGATCAGACT 3626
 Qy 2917 GAGGCTGCCAGCGGTGTACGTGTA 2942
 Db 3627 GAGGCTGCCAGCGGTGTACGTGTA 3652

RESULT 5
 AF207547
 LOCUS AF207547
 DEFINITION Homo sapiens serine/threonine kinase KPM mRNA, linear PRI 01-JUL-2000
 ACCESSION AF207547
 VERSION AF207547.1
 GI:8886763

KEYWORDS
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE
 1 (bases 1 to 4098)
 Hori, T., Takaori-Kondo, A., Kamikubo, Y. and Uchiyama, T.
 AUTHORS Molecular cloning of a novel human protein kinase, kpm, that is
 homologous to warts/lats, a Drosophila tumor suppressor
 TITLE Oncogene 19 (27), 3101-3109 (2000)
 JOURNAL

MEDLINE 20332247
 PUBMED 10871863
 REFERENCE 2 (bases 1 to 4098)
 AUTHORS Hori,T.
 TITLE Direct Submission
 JOURNAL Submitted (20-NOV-1999) Hematology and Oncology, Kyoto University,
 54 Kawaracho, Shogoin, Sakyo, Kyoto, Kyoto Prefecture 606-8507,
 Japan

FEATURES
 source
 1. .4098
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /tissue_type="myeloid cell line"
 375..3641
 /note="similar to Drosophila warts/lats"
 /codon_start=1
 /product="serine/threonine kinase KPM"
 /protein_id="AAF80561.1"
 /db_xref="GI:8886764"
 /translation="MRPKTPATYTSNSRRLQEIREFGLKQPSKSSVQGLPAGPNSD
 TSLDAKVLGSKDARQQOMRATKFGYOKALEIRYSLLPANESGTSAAAEVNRQ
 MLOELVAGCDEMAGRALKOTGSRSEAELEYISKMGYLDPRNEQIVRVKOTSPGK
 GLMPTVTRRPSPEGTGDSFASVHQLSGTPVEGSPFGADGPTALEEMPRVVDYLPFG
 VGHPGCHQHQPCKYGVASVEAAGAHFPLQAGHYGRPHLLVPEPELGYVQVRSFSG
 SKTPEPTGASLPTKGGGPPGAGLAPPPPAAGLYVPHPHKQAGPAAHQLHVGSR
 SQVSPASDPSQSLTPSRNSLNDVLELSSTSVQVQPAATLARDLSQPKLEAPRA
 HVAEPDPCVPVSRNFSNSHPRPCKKAPSLPAPNTVTAHTAAHLHPKSVRL
 RPEPTAVGSPHAWPAPAPAPAPAAEAGLDAKEEHALALGGAGAPFLDVEYGG
 PDRRCPPVPKHLLLRSKSEYDLDSLACAGMEOSLRAGNPEEGSKRSKAGDKG
 GKDKKQIOTSPVYRKNRDEEKRESRIKSYSPYAFKPFMEQHVENVIKTYQVKNRR
 LQLEKNAKGLCEAEQOEKILYQESNRYNRUKRAKMDKSMFEVKIKTLGIFAGEV
 CLACKVDTHALYAMKTRKDKVLRNQVHVKAERDILAEADNEWVKLYSFQDKDS
 LYFYMDIIPGDMNSLIRMEVPEFLAREFYIAELTIAIESVHKMGFTHRIDPNIL
 IDLDGHIKLDGFLGTRFTHNSKYQKSHVRQDSMEPSDLDWDSNRCBGRDLKT
 LEQRARQHORLHSLVGPVNYTAPEVLLKGYTOLCDWMSVGVILFELMVGOPPEL
 APTPTQLKVINWENTLHPIPAQVLSPEARLDLTKLCCSDADRLGRNADDLKAHFF
 FSAIDFSDIRKQAPAPVPTISHPMDTSNFDVDESPFNDAEGSTNADMTLTPNN
 KHPHATYEFTFRFFDDNGYPPFCPKPSGAESAQAESDLSLDLVDQTEGCPQVYV

BASE COUNT 1006 a 1218 c 1130 g 744 t

Query Match 55.48; Score 1746.4; DB 9; Length 4098;
 Best Local Similarity 77.38; Pred. No. 0;
 Matches 2387; Conservative 0; Mismatches 546; Indels 153; Gaps 16;

Qy 1 ATGAGGACACCCGAGCTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATTC 60
 Db 564 ATGAGGACACCCGAGCTTGGACCTTATCAGAAGCCTTGAGGGAATCAGATATTC 623
 Qy 61 CTCCTGCTTTTGGCAACGAGTACGGCATTCTCGGAGCTGCAGAGGTGAACCGGAGATG 120
 Db 624 TTGTTGCTTTTGTGTAATGAATCGGGACCTCTCGAGCTGCAGAGTGAACCGGCAATG 683
 Qy 121 CTTAGGAGTTGGTAATCGGCGATGTACAGGAGATGGCTGGCAGAGCGCTACAGCAT 180
 Db 684 CTGAGGAACCTGGTGAACCGCAGGATGCCACGAGGATGGCTGGCGGAGCTCTCAAGCAG 743
 Qy 181 ACGGCGAGTAGGATGATCGAAGCTGCCPTGGAGTACATCAGTGAAGTGGCTACCTGGAC 240
 Db 744 ACTGGCAGCAGAGCATCGAGCGCGCTCGAGTACATCAGCAGATGGGCTACCTGGAC 803
 Qy 241 CCCAGGAATGACGAGATGTGCGAGTATCAAGCAGACCTCCCGAGAAAGGGCTGGCG 300
 Db 804 CCGAGGAATGACGAGATGTGCGGCTCATTAAGCAGACCTCCCGAGAAAGGGCTCATG 863
 Qy 301 TCCACCGCGTACTCGCGCGCCAGTTTCGAGGCGCAGGGGAAGCACTCCCATCTTAC 360
 Db 864 CCAACCCGAGTGAAGCGGAGCGCAGCTTCGAAGGAACCGGCGATTCGTTTTCGCTCTAC 923
 Qy 361 CACCAGCTGGTGGTGAACACTACGAG-----GGCCCCCGCCCA 399
 Db 924 CACCAGCTGAGCGGTGACCCCTCAGAGGGGCCCAAGCTTCGCGGTGAGCGGCCCAACGCG 983

Qy 400 CTGAGGAGATGCCGCGCAATATTAGACTTTCTCTCCCTGGAGCCGCGGACCGCACC 459
 Db 984 CTGAGGAGATGCCGCGCGCGCTACGTGGACTACCTTTTCCCGGAGTGGCGCCCGCAGGG 1043
 Qy 460 CACGGTGGCCAGGCTACACAGCATCTCCCAAGGGTACAGCACAGCAGTAGAGCAAGT 519
 Db 1044 CCGCGCCACAG---CACCAGCACCACCAAGGGCTACGGTGCCAGCAGTAGAGCAGCA 1100
 Qy 520 GCG-----CACTTCCGGGCGACACACTATGCTGGTGGTCACTACT-----ATCG 564
 Db 1101 GGGGACACATTCCTCCGCTGAGGGCGCGCACTACGGGCGCGCGCACTGCTGGTGGCTGG 1160
 Qy 565 GAGCAGTCTGGGTATGGGTGAGGCGAGTCTCTCTTCCAGAAAGAACAGCGCCACAGAT 624
 Db 1161 GAACCCCTGGGCTACGGAGTGCAGCGACGCCCTCTCTCCAGAGCAAGACGCCCGGAG 1220
 Qy 625 GCC-----TATTCAGCATGCGC---AAGGCCAGGGTGGCCCTCC---CGCCAGCCTC 672
 Db 1221 ACCGGGGTGTACGCCAGCTGCCACGAAGGCCAGGGAGGACCGCCAGGCGCGCGCTC 1280
 Qy 673 ACCTTTCTGCTGCTGGGCTGTACACTGCTCGCACCAAGCCGCGGCTACCCCA 732
 Db 1281 GCTTTCCACACCCCTGCGCGCGGGCTCTACGTGCGCGCACCCACACCAAGCAGCGCGT 1340
 Qy 733 CCTGGGGCCACCCCATTTACATGTTGGGCACCCGGGGTCCCACTTTACTGGCGAAAGC 792
 Db 1341 CCGCGGGCCACCCAGCTGCATGCTGGCTCCCGCAGCAGGTGTTGCCAGCAGCACAGC 1400
 Qy 793 TCTGCACAGGTGCTGGCAGCTGCCAGGCTCCAGGAACAGCCTCAATGCTGCTGTGACAGCTG 852
 Db 1401 CCGCGCAGAGCTGCTCACTCTCCCTCGGGAACAGCCTCAAGCTGATGAATTG 1460
 Qy 853 GG-----CTCCACGGTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
 Db 1461 AGCAGACCTCTCGTCCAGAGTGGCGCGCTGCCACCTTGGCGCGCGGAGTCCCTGCGAG 1520
 Qy 907 AAGAGGCTCTAGAAAGCTTCGCGGC---CGCATGTGGGTGTTTTCGGGCTGGC-----954
 Db 1521 AAGCGGGCTTGGAGCG 1580
 Qy 955 CCGCAGGACCACTCTTCAACAAACCCACAC-----CT 990
 Db 1581 CCGCAGGACCACTCTTCAACAGCAGCAGCAGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1640
 Qy 991 GAGCCTCAGTCCG 1050
 Db 1641 GAGCCTCCTCCCTCG 1700
 Qy 1051 CTTGTGAAGAGCGTGCCTGCTGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1110
 Db 1701 CCGGTGAAGAGCGTGCCTGCTGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1760
 Qy 1111 CCGCCTGGTGGTGGCGC-----CCACAGACCTGTCGCGCTGAG 1149
 Db 1761 CCGCCTGGTGGTGGCG 1820
 Qy 1150 AGCCTGGAGAGCAGGA-----GGCAGCGCAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1191
 Db 1821 GCTTGGAGCGCAGGAGGAGCATGCCCTGGCGTGGGCGCGCGCGCGCGCGCGCGCGCGCG 1880
 Qy 1192 GATGTGGACTATGGCGGCTCCGAGCGAGGTGCCACCGCTTCCCTATTCCTCAAGACATG 1251
 Db 1881 GACGTGAGTACGAGGCG 1940
 Qy 1252 CTGCTGCCCACTAGTCTGAGCAGTACAGCTGGACCTGGACAGCTGTGACACCAAGTGTG 1311
 Db 1941 CTGCTGCCCACTAGTCTGAGCAGTAC-----GACCTGGACAGCTGTGCGCAGCGCATG 1994
 Qy 1312 CAGCAGACTCTGCGAGGGG---CACTGATCTAGACGGGAGTGAACAGCAGCAAAAGGT 1368
 Db 1995 GAGCAGAGCCTTCCCTGGCGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2054

QY 1369 GCGAAGGAGACAAAGCTGGCAGAGACAAAAGCAGATTTCAGACCTCCCGGCTGCCTGTC 1428
 Db 2055 GCCAAGGGGACAAAGCGGAGGATAAAAGCAGATTTCAGACCTCTCCCGTTCCCGTC 2114
 QY 1429 CGCAAGATAGCAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTAT 1488
 Db 2115 CGCAAAACAGCAGAGAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTATC 2174
 QY 1489 GCCTTCAAAATCTTTCAGGAGCAGCAGTGGAGAATGTCAATCAAAACCTACCAAGCAGAG 1548
 Db 2175 GCCTTTAAGTTCTTCATGGAGCAGCAGTGGAGAATGTCAATCAAAACCTACCAAGCAGAG 2234
 QY 1549 GTCAGCGGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGAGCGCGAG 1608
 Db 2235 GTTAACCGGAGGCTGACAGCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGAGCGCTGAG 2294
 QY 1609 CAGCAGCAGATGAGAGAGTCTCTACAGAGAGTCTAACTACAACCGGCTGAAGAGG 1668
 Db 2295 CAGGAGCAGATGCGGAAGATCTCTACAGAGAGTCTAACTACAACAGGTTAAAGAGG 2354
 QY 1669 GCCAAGATGGAACAGTCCATGTTGTGAAATCAAGACTCTAGGATCGGTGCTTTGGG 1728
 Db 2355 GCCAAGATGGAACAGTCTATGTTGTCAAGATCAAAACCTGGGGATCGGTGCTTTGGA 2414
 QY 1729 GAAGTGTGCTGCTGTAACTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGG 1788
 Db 2415 GAAGTGTGCTGCTGTAACTGGAGACTCACGCTCTGTACGCCATGAAGACTCTCAGG 2474
 QY 1789 AAGAGGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 1848
 Db 2475 AAGAGGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTG 2534
 QY 1849 GCTGAAGCAGACATAGTGGGTGGTCAAACTCTACTCTCTTCCTCCAGGACAGAGCAGC 1908
 Db 2535 GCCAGGAGCAGACATAGTGGGTGGTCAAACTCTACTCTCTTCCTCCAGGACAGAGCAGC 2594
 QY 1909 CTGTACTTTGTGATGAGTACATACAGCGGGGATGATGAGCGCTGCTGATCAGGATG 1968
 Db 2595 CTGTACTTTGTGATGAGTACATACCTGCTGGGACATGATGAGCGCTGCTGATCCGGATG 2654
 QY 1969 GAGGTCTTCCCTGAGCAGCTCGCCGCTTCTACATTTGAGAGAGTTGACCTGGCCATTGAA 2028
 Db 2655 GAGGTCTTCCCTGAGCAGCTCGCCGCTTCTACATTCGAGAGAGTCTTTGGCCATTGAG 2714
 QY 2029 AGTGTCCACAGATGGGCTTTATCCACCGGACATCAAGCTGACCACTACTCATGAC 2088
 Db 2715 AGTGTCCACAGATGGGCTTTATCCACCGGACATCAAGCTGACCACTACTCATGAT 2774
 QY 2089 CTGGATGCTCATATTAAGTACAGATTTTGGCCCTCTGCACTGGATTCAGTGGACTCAC 2148
 Db 2775 CTGGATGCTCATATTAAGTACAGATTTTGGCCCTCTGCACTGGATTCAGTGGACTCAC 2834
 QY 2149 AATTCAGATCTACCAAGAGGAAACACATGAGACAGGACAGCATGGAGCCGGTGAC 2208
 Db 2835 AATTCAGATCTACCAAGAGGAAACATGAGACAGGACAGCATGGAGCCGGTGAC 2894
 QY 2209 CTCTGGAGGATGTTTCAACTGTCTGCTGGAGACAGGTTAAAGACCTCGAGCAGAGG 2268
 Db 2895 CTCTGGAGGATGTTTCAACTGTCTGCTGGAGACAGGTTAAAGACCTCGAGCAGAGG 2954
 QY 2269 CGCGAAGACAGCAGCAGAGGCTGCGCATCTCTGTGCGGACACCAAAATACATC 2328
 Db 2955 CGCGAAGACAGCAGCAGAGGCTGCGCATCTCTGTGCGGACACCAAAATACATC 3014
 QY 2329 GCTCCGGAGGCTCTTCCGCAAGAGGTACAGCAGCTGTGACTGTGGAGCGCTCGGT 2388
 Db 3015 GCAACCGAGGCTCTCTCCGCAAGAGGTACACTCAACTCTGTGACTGTGGAGTGTGA 3074
 QY 2389 GTGATTTCTTTGAGATGCTGTGGAGCGGCTTTCTTTGGCCCCCACCACACAGAG 2448
 Db 3075 GTGATTTCTTTGAGATGCTGTGGAGCGGCTTTCTTTGGCCCCCACCACACAGAA 3134
 QY 2449 ACGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGCATATCCCTACGAGGTGAGGCTC 2508

Db 3135 ACCCAGCTGAAGTGTATCAACTGGGAGACACGCTCCACATTCAGCCCGCAGGTGAAGCTG 3194
 QY 2509 AGCGCTGAGGCGGAGACCTCATCAGAGAGCTGTGCTGCGCGCTGACTGCGCCTGGGC 2568
 Db 3195 AGCCCTGAGGCGGAGGACCTCATCACCAGCTGTGCTGCTCCGAGACACCGCCTGGGG 3254
 QY 2569 AGGATGSGGAGATGACCTCAAGGCACACCGCTTCTTCAACACCATCGACTTTTCCCT 2628
 Db 3255 CGGAATGGGCGGATGACCTGAAGGCCACCCCTTCTTCAAGGCCATGACTTCTCCAT 3314
 QY 2629 GACATCCGAAAGCAGGCTGCACCTACCTCCACCATCAGCCACCCATGAGACACTCC 2688
 Db 3315 GACATCCGAAAGCAGGACGCGCCCTACCTTCCACCATCAGCCACCCATGAGACACTCG 3374
 QY 2689 AATTTGACCCGCTGATGAAGAAACCCCTTGCAGAGCGCCGAGGAGCGGCAAG 2748
 Db 3375 AATTTGACCCGCTGATGAAGAAACCCCTTGCAGAGCGCCGAGGAGGAGCGGCAAG 3434
 QY 2749 GCTGGGACACGCTGCGCTCCCGCAGCAGCAAGCATCCAGAGCAGCCCTTCTATGATTC 2808
 Db 3435 GCTGGGACACATCAGCTCGCCCATTAACAGCATCTTGCAGCAGCATTTTACGAATTC 3494
 QY 2809 ACCTTCGCGAGTCTTCGATGACAGCGCTATCCCTTCGGTGGCGGAGCCCTCAG-- 2866
 Db 3495 ACCTTCGCGAGTCTTCGATGACAGCGCTATCCCTTCGGTGGCGGAGCCCTCAGGA 3554
 QY 2867 -----AGCCCGCAGAGAGTGCAGACCCAGGGGATGCGGACTTGGAGGCTGGCGGC 2916
 Db 3555 GCAGAGCTTTCAGAGCTGAGAGCTCAGATTAGAAGCTCTGATCTGGTGGATCAGACT 3614
 QY 2917 GAGGCTGCCAGCGGCTGACGTGTA 2942
 Db 3615 GAAGGCTGCCAGCGCTGATGTA 3640

RESULT 6
 AB028019 3412 bp mRNA linear PRI 09-MAR-2000.
 LOCUS
 DEFINITION Homo sapiens HSLATS2 mRNA for large tumor suppressor 2, partial cds.

ACCESSION AB028019
 VERSION AB028019.1 GI:7212789
 KEYWORDS large tumor suppressor 2; HsLats2.
 SOURCE Homo sapiens cDNA to mRNA.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.

REFERENCE 1 (sites)
 Yabuta, N., Fujii, T., Copeland, N.G., Gilbert, D.J., Jenkins, N.A., Nishiguchi, H., Endo, Y., Toji, S., Tanaka, H., Nishimune, Y. and Nojima, H.
 Structure, expression, and chromosome mapping of LATS2, a mammalian homologue of the Drosophila tumor suppressor gene lats/warts
 Genomics 63 (2), 263-270 (2000)
 20139436
 2 (bases 1 to 3412)
 Nojima, H. and Nishiguchi, H.
 Direct Submission
 Submitted (27-MAY-1999), Hiroshi Nojima, Research Institute for Microbial Diseases, Osaka University, Department of Molecular Genetics, Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
 (E-mail: nojima@biken.osaka-u.ac.jp, Tel: 81-6-6875-3980, Fax: 81-6-6875-5192)

FEATURES
 source
 1. 3412 Location/Qualifiers
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 1. 3412
 /gene="HsLATS2"
 <1. 3143
 /gene="HsLATS2"
 /codon_start=3


```

/product="large tumor suppressor 2"
/db_xref="gi:7212790"
/translation="NSDTSLEKADKSGKDATSSKQMRATPKFGPYOKALREIRYSL
PFANESGTSAAEYNROMLOELVNAVGDQEMAGRALKOTGTSERSEALEALYSKGYLD
PRNEOLIVRIKQTSFGKGLMPTVTRRPSFEGTSDSFASYSYHQLSGTPEYGFSGADG
TALSEMRPVIYDIFPGVPHGPHQHPKPGYASVAAAGAHFPLQAGHYRPHLL
VPGLEPLGVORSFOSKTPETGGYASLPTKGGGPPGAGLAFPPPAAGLYVPHPL
HKQAPVAHQHVLGSRQSFASPPQSLTTPSRNSLNVLDYELSTSVQVQPAATL
ARRDSLOKPLGLEAPRAHVPDPVPSRTNSFNHSHQPPGPKAPPSLUPAPNTVT
AVSAHILHPVKSVRVLRPEPQTAVGSPSHPAWVAPAPAPAPAPAPAEGLDAKEHA
LALGAGAPLDVVEYGGPDRSARLPTSRSCCAKSEQYDLDSLCAQBSQSLRAGNE
PEGDKSRKSKAGDKNGKQKQIOTSPVVRKNSRDEKRESRIKSYSPYAFKRFMEQ
HVENIKTYQOKVNRLOEQEMAKAGLCEAEQOMRKILYKESNINRLKRAKMDK
NEWVKLYSFQDGLSEVCLACKVDTHALYAMKLRKVDLNRNOVAHVKAERDILAEAD
HMKGFIHRDILDLGHIKILDTGFLCTGFRWTHNSKYKOGSHVRQDSPEPSD
LWDDVSNCRGDRKLTLEQARKOHRSLAHSVLTNYTAPEVLLRKYITQLCDWNS
VGVILFEMLQGPFLAPTETETQLKINMENTUHI PAQVKLSPEARLDLITKLCSAD
HRLGRNADDLKAHPFTSAIDFSDIRKQAPYVPTISHPMKTSNFPDVEESPWDA
SEGSKAMDTLTSNNKHPEHAFYEFTFRFDDNGVYFPRCPKPSGAESAQESDLE
SSDLVDOTEGQPVV"
3364..3369
/polya_signal
BASE COUNT 801 a 1066 c 975 g 570 t
/gene="HsLATS2"

```

```

Query Match 54.4%; Score 1717.2; DB 9; Length 3412;
Best Local Similarity 77.2%; Pred. No. 0;
Matches 2382; Conservative 0; Mismatches 548; Indels 156; Gaps 18;

Qy 1 ATGAGAGCACCCGAGTGTGGACCTTATCAAAAGCTCTCAGGAAATCCGATATATCC 60
Db 69 ATGAGAGCACCCGAGTGTGGACCTTATCAGAAAGCTTGGAGGAAATCAGATATATCC 128

Qy 61 CTCCTGCTTTTGGCAAGAGTCAGGACCTTCGGCAGCTTCGAGAGGTGAACGGGAGATG 120
Db 129 TGTGCTGCTTTTGGCAAGAGTCAGGACCTTCGAGCTTCGAGAGGTGAACGGGAAATG 188

Qy 121 CTCAGAGAGTGTGAATTCGGCAGTGTGACGAGAGAGTGTGGCAGAGCTTCACGACG 180
Db 189 CTCAGAGAGTGTGAATTCGGCAGTGTGACGAGAGAGTGTGGCAGAGCTTCACGACG 248

Qy 191 ACGGCACTAGAGTATCGAAGTGCCTTGGAGTACATCAGTAAGATGGGTACTCTGAC 240
Db 249 ACTGCAAGAGAGTATCGAAGTGCCTTGGAGTACATCAGTAAGATGGGTACTCTGAC 308

Qy 241 CCCAGGATGAGCAGTGTGGAGTATCAGCAGAGCTCCCGAGGAAAGGCGCTGGCG 300
Db 309 CCGAGGATGAGCAGTGTGGAGTATCAGCAGAGCTCCCGAGGAAAGGCGCTCATG 368

Qy 301 TCCACCCCGGTGACTTCGGCGGCCAGTTCGAGGACGAGGAGGAGTCTCCCATCTAC 360
Db 369 CCAACCCCGGTGACTTCGGCGGCCAGTTCGAGGACGAGGAGGAGTCTCCCATCTAC 428

Qy 361 CACAGCTGGGTGTGCAAACTACAG-----GGCCCGCGCCCA 399
Db 429 CACAGCTGGGTGTGCAAACTACAG-----GGCCCGCGCCCA 488

Qy 400 CTGAGGAGATGCGCGGCAATATTAGACTTCTCTCCCTGGAGCGCGGAGCGGAC 459
Db 489 CTGAGGAGATGCGCGGCGGCTAGTGGACTTACCTTTTCCCGGAGTGGCGCCCGCC 548

Qy 460 CACGGTGCAGGCTCACAGCATCTCCCAAGGGTACAGCAGCAGTAGAGCCAGT 519
Db 549 CCGGGCACAG-----CACAGCAGCAGCAGCAGCAGCAGCAGTAGAGCGAGCA 605

Qy 520 GCG-----CACTTTCGGGCGACACACTATGCTGTGCTCATCTACT-----ATCG 564
Db 606 GGGGCACACTTTCGGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGGCGG 665

Qy 565 GAGCAGTCTGGGTATGGGTGCAGGAGTTCCTCTCTCCAGAACAGAGCCACAGAT 624
Db 1737 GTTAAACCGGAGCTGCAGTGGAGCAAGAAATGGCCAAAGCTGTGTGAAGCTGAG 1796.

```

```

Db 666 GAACCCCTGGGCTACGAGTGCAGCGACCCCTCTCTCCAGAGCAAGACGCCGCCGAG 725
Qy 625 GCG-----TATTCAGCATGCG-----AAGGCCAGGCTGGCCCTCC-----CGCCAGCTC 672
Db 726 ACCGGGGTTACCGCAGCGCTGCCACGAGAGGCGCAGGAGGAGGAGGCGCGGCGCTC 785
Qy 673 ACCTTCTCTGCGCATGCTGGGCTGTACAGTCTCGCAGCACAAGCGCGGCTTACCCCA 732
Db 786 GCTTTCCCAACCCCTGCGCGCGGCTCTACGTGCGCGCACCAACACACAGCAGCGCT 845
Qy 733 CTGGGGGCCACCCATTCATATGTGTGGACACCCCGGGTCCACGCTTTTACTGGCGAAGC 792
Db 846 CCGCTGGGCCACCACTGATGCTGTGGCTCCCGCAGCAGGCTGTCCGCGCAGCAGCAGC 905
Qy 793 TCTGCACAGGCTGTCTGCGACCGCTGCAGTCTGCAGTCCACCTGGCAGCGGAGCTTCGCTGAG 852
Db 906 CCGCGCAGAGGCTGCTCACTCCCTCGCGAAGACCTCAACGTGGAGCTGTATGAATG 965
Qy 853 GG-----CTCCACGGTGCCTGTCTGAGTCTGCAGTCCACCTGGCAGCGGAGCTTCGCTGAG 906
Db 966 AGCAGCAGCTCGTCCAGCAGTGGCGGCTGCCACCTTGGCGCGGAGCTTCCTGCGAG 1025
Qy 907 AAGCAGGCTTAGAAGCTTCGGGCG-----CGCATGTGGCTTTTTCGGGCTGGC----- 954
Db 1026 AAGCGGGCTGGAGGCGCGCGCGCGCAGCTGGGCTTCCGGCCTGACTGCCAGTG 1085
Qy 955 CCCAGCAGGACCACTCTTCAACCAACCCACAAC-----CT 990
Db 1086 CCCAGCAGGACCACTCTTCAACAGCAGCAGCAGCGCGCGCGCTTCGCGCTGGCAAGGCC 1145
Qy 991 GAGCCTCACTGCGCGCGCGCGCGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCGCG 1050
Db 1146 GAGCCTCCCTGCGCGCGCGCGCGCGCGCTGAGCGCGCGCGCGCGCGCGCGCGCGCG 1205
Qy 1051 CTGTGAAGAGCTGTGCTGTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1110
Db 1206 CCGTGAAGAGCTGTGCTGTGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1265
Qy 1111 CCGCTGCGGTGGCTGCGC-----CCACAGCAGCTGCCACTGAG 1149
Db 1266 CCGCTGCGGTGGCTGCGC-----CCCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1325
Qy 1150 AGCCTGGAGAGCAAGGA-----GGCAGCGCGAGCGCGCGCGCGCGCGCGCGCG 1191
Db 1326 GCGTGGAGCGCAAGGAGGAGTGCCTGCGCTGGCGCGCGCGCGCGCGCGCGCGCGCG 1385
Qy 1192 GATGTGAGTATGGGGGCTCGCGGCGAGGTGCCACCGCTTCGCTATCAAGAGCTTG 1251
Db 1386 GACGTGGAGTACGAGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1444
Qy 1252 CTGCTGCCAGTAACTGTGAGCAGTACAGCGTGGACCTGGACAGCTGTGCGACCGAGTGTG 1311
Db 1445 CTGCTGCCA--AGTGGAGCAGTAC-----GACCTGGAGCAGCTGTGCGCGAGCGATG 1496
Qy 1312 CAGCAGAGTCTGCGAGGGG-----CACTGATCTAGACGGGAGTGACAGAGCGCACAAAGGT 1368
Db 1497 GAGCAGAGCTCGCTGGGGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1556
Qy 1369 GCGAAGGAGACAAAGCTGGCGAGAGCAAAAGCAGATTCAGACCTTCCCGGCTGCTGTG 1428
Db 1557 GCCAAGGGGACAAAGGCGGAAAGGATTAAGAGCAGATTCAGACCTTCTCCGCTTCCGCTC 1616
Qy 1429 CGCAGGATACGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTCTACTCCCTTAT 1488
Db 1617 CGCAAAACAGCAGAGAGAGAGAGAGAGAGTACGCGATCAAGAGCTACTCGCCATAC 1676
Qy 1489 GCCTTCAATTTCTTATGGAGCAACAGTGGAGATGTCAATAAACTTACACAGCAAG 1548
Db 1677 GCCTTTAGTCTTATGGAGCAGCAGCTGGAGATGTCAATAAACTTACACAGCAAG 1736
Qy 1549 GTCAGCGGAGGCTTACAGTGGAGCAGGAAATGCCAAAGCTGGGCTCTGTGAGCGCGAG 1608
Db 1737 GTTAAACCGGAGCTGCAGTGGAGCAAGAAATGGCCAAAGCTGTGTGAAGCTGAG 1796.

```

QY 1609 CAGGAGCAGATGAGGAGATCTCTTACCAGAGGAGCTTAACCTACAAACGGCTGAGAGG 1668
DB 1797 CAGGAGCAGATGCGGAGATCTCTTACCAGAGGAGCTTAACCTACAAACGGCTGAGAGG 1856
QY 1669 GCCAAGATGACAAAGTCCATCTTTGTGAATCAAGACTCTAGGATCGGCTGGTGG 1728
DB 1857 GCCAAGATGACAAAGTCTTATGTTGTGAAGATCAAAACCTGGGATCGGCTGGTGG 1916
QY 1729 GAAGTGTGCTGCTGCTGTAAGCTGACACTCAAGCTGCTGACGCTGCTGATCAGGATG 1788
DB 1917 GAAGTGTGCTGCTGCTGTAAGCTGACACTCAAGCTGCTGACGCTGCTGATCAGGATG 1976
QY 1789 AGAAGGATGCTGCTGTAAGCTGACACTCAAGCTGCTGACGCTGCTGATCAGGATG 1848
DB 1977 AGAAGGATGCTGCTGTAAGCTGACACTCAAGCTGCTGACGCTGCTGATCAGGATG 2036
QY 1849 GCTGAAGCAGACAAAGTGGGTGCTCAAGCTGCTGACGCTGCTGATCAGGATG 1908
DB 2037 GCTGAAGCAGACAAAGTGGGTGCTCAAGCTGCTGACGCTGCTGATCAGGATG 2096
QY 1909 CTGTACTTTGTGATGACTACATACAGGCGGGGATATGATGAGCTGCTGATCAGGATG 1968
DB 2097 CTGTACTTTGTGATGACTACATACAGGCGGGGATATGATGAGCTGCTGATCAGGATG 2156
QY 1969 GAGGTCTTCCCTGACACTGCGCGCTTACATGTCAGAGTGAACCTGGCCATTGAA 2028
DB 2157 GAGGTCTTCCCTGACACTGCGCGCTTACATGTCAGAGTGAACCTGGCCATTGAA 2216
QY 2029 AGTGTCCCAAGATGGCTTTATCCACCGGACATCAAGCTGACATCACTATCATGAC 2088
DB 2217 AGTGTCCCAAGATGGCTTTATCCACCGGACATCAAGCTGACATCACTATCATGAC 2276
QY 2089 CTGTGATGCTATATTAAGCTGACAGATTTTGGCTCTGCACTGATGACGCTGACATC 2148
DB 2277 CTGTGATGCTATATTAAGCTGACAGATTTTGGCTCTGCACTGATGACGCTGACATC 2336
QY 2149 AATTCCAAAGTACTACAGAGGAGGACATGATGACAGACAGCATGGAGCCCGGTGAC 2208
DB 2337 AATTCCAAAGTACTACAGAGGAGGACATGATGACAGACAGCATGGAGCCCGGTGAC 2396
QY 2209 CTCTGGGACGATTTTCCAACTGCTGCTGGAGAGGTTAAAGACCTTGGAGCAGAGG 2268
DB 2397 CTCTGGGACGATTTTCCAACTGCTGCTGGAGAGGTTAAAGACCTTGGAGCAGAGG 2456
QY 2269 GCGCAGAGCAGCAGAGGCTGCTGGACATCTCTTGTGGGACACCAATATCATC 2328
DB 2457 GCGCAGAGCAGCAGAGGCTGCTGGACATCTCTTGTGGGACACCAATATCATC 2516
QY 2329 GCTCCGAGGCTGCTTCTCCGCAAGGCTACACGAGCTCTGACGCTGCTGAGCCTGCT 2388
DB 2517 GCTCCGAGGCTGCTTCTCCGCAAGGCTACACGAGCTCTGACGCTGCTGAGCCTGCT 2576
QY 2389 GTGATCTCTTTGAGATGCTGCTGGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2448
DB 2577 GTGATCTCTTTGAGATGCTGCTGGGACGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2636
QY 2449 ACCGAGTGAAGTGAATCACTGCGAGAGCAGCCTGCTGCTGCTGCTGCTGCTGCTG 2508
DB 2637 ACCGAGTGAAGTGAATCACTGCGAGAGCAGCCTGCTGCTGCTGCTGCTGCTGCTG 2696
QY 2509 AGCGCTGAGCCCGGAGACTCTATCAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2568
DB 2697 AGCGCTGAGCCCGGAGACTCTATCAGAAAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2756
QY 2569 AGGATGAGGCGAGATGACCTCAAGGACACCGCTTCTTCAACACCATCATCTTTCCCGT 2628
DB 2757 AGGATGAGGCGAGATGACCTCAAGGACACCGCTTCTTCAACACCATCATCTTTCCCGT 2816
QY 2629 GACATCCGAAGAGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2688
DB 2817 GACATCCGAAGAGGCTGACCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2876

QY 2689 AATTTTGACCGCTGGATGAAGAAAGCCCTTGGCACAGGCGCCAGCGGAGAGAGCGCAAG 2748
DB 2877 AATTTTGACCGCTGGATGAAGAAAGCCCTTGGAACAGTGCAGGAAGGTAGCACCAAG 2936
QY 2749 GCCTGGGACACGCTGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTC 2808
DB 2937 GCCTGGGACACGCTGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTC 2996
QY 2809 ACCTTCCGACAGTCTTGGATGACAAAGGCTATCCCTTCCGGTCCCGGAGCGCTCAG-- 2866
DB 2997 ACCTTCCGACAGTCTTGGATGACAAAGGCTATCCCTTCCGGTCCCGGAGCGCTCAG-- 3056
QY 2867 -----AGCCCGCAGAGAGTGCAGACCGGAGTGGGACTTGGAGAGTGGCGCC 2916
DB 3057 GCAGAGCTTCCAGAGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGGTGGATCAGACT 3116
QY 2917 GAGGCTGCCAGCGCTGCTACGTGTA 2942
DB 3117 GAGGCTGCCAGCGCTGCTACGTGTA 3142

RESULT 7
AR201456
LOCUS
DEFINITION Sequence 5 from patent US 6359193.
ACCESSION AR201456
VERSION AR201456.1 GI:20252344
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1. (bases 1 to 3213)
AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 5 19-MAR-2002;
FEATURES
Location/Qualifiers
1..3213
/organism="unknown"
BASE COUNT 946 a 764 c 696 g 807 t
ORIGIN

Query Match 25.8%; Score 812.6; DB 6; Length 3213;
Best Local Similarity 65.1%; Pred. No. 5.5e-170;
Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

QY 985 CAACCTGAGCGCTCTACTGCGCGCCCGCCACACAGGTCACCGCGTGCAGCGCCGACACATC 1044
DB 919 CACTCTGCTAATCTCAGCGCTTCTGCCACTACAGTCACTGCCATCACACCGCTCTCTATT 978
QY 1045 CTTCAACCTGTGAAGAGCGTGGTGTGCTGCGCGCCCGAGCGCCACAGACCGCTGGGCCC 1104
DB 979 CAACAGCGCGTCAAAAGCATGCGCGTCTTGAAACACAGAGTGCAGACTGCTTTAGCCCA 1038
QY 1105 TCGACACCGCGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 1164
DB 1039 ACCCTATCTCTTGGATGCCAGAGTTCAGACTGTTGAGCGCTTACCCCTTTTCTCAG 1098
QY 1165 GAGGCGAGCGAGCGCGCCACACCGCTGGATGTGGACTATGGGCTCGAGCGAGGTGC 1224
DB 1099 GGTACAGCTTCAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1158
QY 1225 CCACCGCTTCCGATPCCAAAGCACTTGTGCTGCGCCAGTAAGTCTGAGCAGTACAGCGTG 1284
DB 1159 CCACCGCGCTTATCCAAACATCTGCTACACCAAAACCCATCTGCTCCCTCCATATGAG 1218
QY 1285 GACCTGGACAGCGCTGCGCAGAGTGCAGCAGAGTGCAGGAGGCGGCGCTGATCTAGAC 1344
DB 1219 TCAGTAAGTAAAGCTGCGCAGAGTGCAGCAGAGTGCAGCAGAGTGCAGCAGAGTGCAG 1278
QY 1345 GGGAGTGAACAGAGCCACAAAGGTGCGAAGGAGCAGAAAGCTGGCAGAGCAAAAGCAG 1404
DB 1279 AAGAGTGGGACAGT-----GGTCACTCTGGGATAAAGAAAGAAACAG 1323

QY 1405 ATTACAGACCTCCCGGTGCTCTCCGCAAGAATAGCAGAGATGAAGAGAGAGAGAGTCT 1464
Db 1324 ATTACAACTTCCCTATCATCTGTCGGAAGAAACAGAAAGATGAAGAGAGAGAGTCT 1383
QY 1465 CCATCAAGAGTACTCCCTTATCCCTTCAAAATCTTCTCATGAGCAACACGCTGGAGAT 1524
Db 1384 CGATTCAGAGTACTCCCAAGGCTTTAAGTCTTCTCATGAGCAGCAGCTAGAGAAC 1443
QY 1525 GTCATCAAAACCTACACAGAGAGTTCAGCCGAGGCTACAGCTGGAGCAGGAAATGGCC 1584
Db 1444 GTCCTGAAGTCTCATCAGCAGGCTGTCATCGAAGAACAGCTAGAAATGAATGATG 1503
QY 1585 AAGCTGGGCTCTGTGAGCGCAGCAGAGAGATGAGGAAGATCTCTACCAAGAGAG 1644
Db 1504 CGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAGATGCTTTGCCAGAAAGAG 1563
QY 1645 TCTAAGTACACCGCTGAAGAGGCCAAGATGGACAGTCTATGTTGTGAATCAAG 1704
Db 1564 TCTAAGTATATTCGCTTTAAAGGGCTAAATGGACAGTCTATGTTGTGAAGATAAG 1623
QY 1705 ACTCTAGGCTCGGTGCTTTGGGAAGTGTGCTGCTTGTAAAGCTGGACACTCAGCT 1764
Db 1624 ACATTAAGTAAGAGCGTTTGGTGAAGTCTCTAGCAAGAAAGTCAATCAAGCT 1683
QY 1765 CTGTAGGCTGAAGACTCTCAGGAAGAGAGTGTCTTGAACCGGAATCAAGTGGCCAT 1824
Db 1684 TTGTATGCAACAAAGACTCTTCCGAAAGAGAGCTTCTGCTCCGAAATCAGGTGCTCAT 1743
QY 1825 GTCAGGCTGAGGAGACATCTGCTGAAGCAGACATGAGTGGTGGTCAAACTCTAC 1884
Db 1744 GTGAAGCGGAGAGGATATCTAGCAGAGCGGACATGAGTGGTGGTCCGCTGTAC 1803
QY 1885 TACTCTTCCAGCAAGCAGCAGCTGTACTTGTGATGAGTACATACCGCGGGAT 1944
Db 1804 TACTCTTCCAGCAAGCAGCAGCTGTACTTGTGATGAGTACATCTCTGGGGGAT 1863
QY 1945 ATGATGAGCTCTGATCAGGATGAGGTCTTCCCTGAGCAGCTCCGCGCTTCTACAT 2004
Db 1864 ATGATGAGCTTATTAATAGATGGCATCTTCTCGTGAATCTGCGAGGATCTACATA 1923
QY 2005 GCAGAGTGTACCTGGCCATCAAGATGTCACAGATGGCTTTTATCCACCGGACATC 2064
Db 1924 GCAGAGTGTACCTGGCAGTGAAGATGTCATAAATGGTGTATTCATAGAGATAT 1983
QY 2065 AAGCTCTCAACATATCTATCAGCTGGATGCTATTAATGATGACAGATTTGGCCCTC 2124
Db 1984 AAGCTCTCAACATATTTGATGACCGTGTATGCGCATATTAATGATGACATTTGGCTTG 2043
QY 2125 TCACTGGATCAGTGGACTCACAATTCACAGTACTACAGAAAGGAGCAACATGAGA 2184
Db 2044 TCACTGGCTTCAGATGGACATGACTCAGAGTACTACAGAGTGGGATCACCACCG 2103
QY 2185 CAGGACAGCATGGACCGGTGACCTCTGGGAGATGTTTCACTGCTGCTGGAGAC 2244
Db 2104 CAAGATGATGATTTTCAGTAAAGATGGGAGATGCTTCAATGCTGGTGGGAGAC 2163
QY 2245 AGTTTAAAGACCTCGAGCAGAGGCGCAAGAGCAGCAGCAGCTGCTGCGCATTTCT 2304
Db 2164 AGACTGAAGCCTAGGAGGAGAGCTGCTCCGAGCAGCAGCTGCTACCCATCT 2223
QY 2305 CTGTGGGACACCAATTTACATCCTCGGAGGTGCTTCCGCAAGAGGTACAGCAG 2364
Db 2224 CTGGTTGGGACTCCCAATATATTCACCTGAAGTCTACTGCAAGCAGATATACAG 2283
QY 2365 CTCTGACTGTGGAGCTCGGTGATCTCTTGTAGATCTGTTGGGAGCGGCT 2424
Db 2284 CTGTGTGACTGTGGAGTGTGGTGTATTTCTTGTGAATGTTGGTGGGCAAACTCT 2343
QY 2425 TTCTTGGGCCCCCCCCACAGCAGCAGCAGCTGAGGTGATCACTGAGGAGCAGCTG 2484
Db 2344 TTCTTGGCACAACCCCATTAACACACAAATGAAGTTATCATCTGGCAAACTCTCTA 2403
QY 2485 CATATCCCTACGAGGTGAGGCTCAGCGCTGAGCGCCGAGACCTCATCAGAAAGTGTGC 2544

Db 2404 CACATCCCTCTCAAGCTAAGCTGAGTCTGAAGCTCTGACCTCATTTATCAACTGTGT 2463
QY 2545 TCGCGGCTGACTCGCGCTGGCAGGATGGGAGATGACTCAAGCAGCAGCCGTTTC 2604
Db 2464 CGAGGACCAAGAGACCGCTCGGCAAGACGGTGTGTGAGATAAAGCTCATCATTT 2523
QY 2605 TTCAACACCATCGACTTTTCCGCTGACATCCGAAAGCAGGCTGCACCTTACCTCCAC 2664
Db 2524 TTTAAGACCATCGATTTCTTAGTATCTGACAGCAGCTGCTTCTACATCCCTAA 2583
QY 2665 ATCAGCCACCCATGGACACCTCCATTTTGTACCCGCTGGATGAAGAAAGCCCTGGC 2724
Db 2584 ATCAGCATCCCAACAGATACATCCCAATTTGACCCCTGTTGATCCTGATAAATG 2643
QY 2725 GAGGACGCGGAGAG---AGCCCAAGGCTGGGACACGCTGGCTCCCGCAGCAGCAG 2781
Db 2644 GATGCGAGGAGGAGAAATATCAGTGACACTCTGAGCGGATGTTATAAATGGGAG 2703
QY 2782 CATCAGACGACGCTCTCTATGATGTTTCCGAGGTTCTTCTGATCAGACAGGCTAT 2841
Db 2704 CACCCGAGCAGCTTCTCTATGATGTTTCCGAGGTTTCTGATCAGACATGGCTAC 2763
QY 2842 CCTTCCGCTGCCGAAAGCCCTCAGAGCCGCGCAGAGAGT 2880
Db 2764 CCATATATTATCCAAAGCCTATTGAGTATGAATACATT 2802

RESULT 8
AF104414
LOCUS AF104414 3213 bp mRNA linear ROD 18-MAR-1999
DEFINITION Mus musculus large tumor suppressor 1 (Lats1) mRNA, partial cds.
ACCESSION AF104414
VERSION AF104414.1 GI:4324435
KEYWORDS
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 3213)
AUTHORS Xu, T., Wang, W., Zhang, S., Stewart, R.A. and Yu, W.
TITLE Identifying tumor suppressors in genetic mosaics: the Drosophila lats gene encodes a putative protein kinase
JOURNAL Development 121 (4), 1053-1063 (1995)
MEDLINE 95262551
PUBMED 7743921
REFERENCE 2 (bases 1 to 3213)
AUTHORS Tao, W., Zhang, S., Turenchalk, G.S., Stewart, R.A., St John, M.A., Chen, W. and Xu, T.
TITLE Human homologue of the Drosophila melanogaster lats tumour suppressor modulates CDC2 activity
JOURNAL Nat. Genet. 21 (2), 177-181 (1999)
MEDLINE 99440767
PUBMED 9988268
REFERENCE 3 (bases 1 to 3213)
AUTHORS St John, M.A., Tao, W., Fei, X., Fukumoto, R., Carcangiu, M.L., Brownstein, D.G., Parlow, A.F., McGrath, J. and Xu, T.
TITLE Mice deficient of Lats1 develop soft-tissue sarcomas, ovarian tumours and pituitary dysfunction
JOURNAL Nat. Genet. 21 (2), 182-186 (1999)
MEDLINE 99140768
PUBMED 9988269
REFERENCE 4 (bases 1 to 3213)
AUTHORS Tao, W. and Xu, T.
TITLE Direct Submission
JOURNAL Submitted (04-NOV-1998) Genetics, HHMI, Yale University, 295 Congress Ave BCM 254D, New Haven, CT 06536, USA
FEATURES
Location/Qualifiers
1. 3213
/organism="Mus musculus"
/db_xref="taxon:10090"
/tissue_type="brain"
/dev_stage="newborn"

Query Match	25.88;	Score 812.6;	DB 10;	Length 3213;
Best Local Similarity	65.18;	Pred. No. 5.5e-170;		
Matches 1237;	Conservative 0;	Mismatches 644;	Indels 18;	Gaps 2;
Qy	985	CAACCTGAGCCCTCTACTGCCCGCCCCCAACAGGTACCCCTGTAGCGCCGACACATC	1044	
Db	919	CACCTGCTAAATTCAGCCCTTCTGCCACTACAGTCTACTGCCATCACACCCGCTCTATT	978	
Qy	1045	CTTACCCCTGTGAAGAGCGTGGTGTGTGCGGCCGAGCCGACAGACGCGTGGGGCCC	1104	
Db	979	CACAGCCCGTGAAGCATGCGCGTCTGAAACACAGAGCTGCAGACTGCTTTAGCCCA	1038	
Qy	1105	TCGCAACCCCGCTGGTGGCTGCGCCACAGACCTGCCACTGAGAGCCTGGAGACGAAG	1164	
Db	1039	ACCATCTCTTTGGATGCGCAGCCAGTTCAGACTGTTGACCTACCCCTTTTCTGAG	1098	
Qy	1165	GAGGGCAGCGACGCCACACCCGCTGGATGTGGACTATGGCGCTCCGAGCCAGGTGC	1224	
Db	1099	GGTACAGCTTCAAGTGTGCTGTCATCCACCTGTTGCTGAAGCTCCAAGCTATCAAGGT	1158	
Qy	1225	CCACGGCTCCGTATCCAAAGCAGCTGTGCTGCCAGTAAGTCTGAGCAGTACACGCTG	1284	
Db	1159	CCACACCGGCTTTATCCAAACATCTGTACACAAACCACATCTGTCCCTCCATATGAG	1218	
Qy	1285	GACCTGCACACCTGTGCACAGTGTGCAGAGAGTCTGCAGGGGGCAGCTGATCTAGAC	1344	
Db	1219	TCAGTAAGTAAGCCCTGCAAGATGAACACCTAGCTTACCACGAAGAGATGATGTGAG	1278	
Qy	1345	GGGAGTGCACAGACCCACAAAGGTGCGAAGGGAGACAAAGCTGGCAGACAGACAAAAGCAG	1404	
Db	1279	AAGAGTGGGACAGT-----GGTGACTCTGGGATATAAGAAACAGAAACAG	1323	
Qy	1405	ATTACAGACTCCCGGTGCTGTCCGCAAGATAGCAGAGATGAAGACAGAGAGTCT	1464	
Db	1324	ATTACAACTTCACCTATCACTGTTCGGAACAAACAGAAAGATGAAGACGAAGAGAGTCT	1383	
Qy	1465	GCATCAGAGTTACTCCCTTTATGCTTTCAAAATTTCTCATGGACCAACAGTGGAGAAAT	1524	
Db	1384	CGGATTCAGAGTTACTCCCAACAGGCCCTTTAAGTTCTTCATCGGACGACGTAGAGAAC	1443	
Qy	1525	GTCAATCAAAACCTTACCAGCAGAGGTGACGCCGAGCCTACGCTGGACGAGAAATGCC	1584	
Db	1444	GTCTCGAAGTCTCATACGACGCGTCTGCATTCGGGAAGACGCTGAAATGAATGATG	1503	

```
Db 2584 ATCACGCATCCACACAGATACATCCATTCGACCCCTGTGTATCTCTGATAAATTTGGAGC 2643
Qy 2725 GAGGCCAGCGGAGAG---AGGCCAGGCGCTGGGACACACCTGCCCTCCCGCCAGCAGAG 2781
Db 2644 GATGGCAGCGAGGAGAAATATCAGTGACACTTGAGCGGATGGGTATATAAATGGGAAG 2703
Qy 2782 CATCCAGACAGCGCTTCTATGAGTTCACTTCCGAGCTTCTTCGATGACACAGGCTAT 2841
Db 2704 CACCCGAGCAGCGCTTCTATGAGTTCACTTCCGAGGCTTTCGATGACATGGCTAC 2763
Qy 2842 CCCTCCGCTGGCCGAGCCCTCAGAGCCCGCAGAGAGT 2880
Db 2764 CCATATAATTATCAAGCGCTATTGAGTATGAATACATT 2802

RESULT 9
AR201455
LOCUS AR201455 3984 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 3 from patent US 6359193.
ACCESSION AR201455
VERSION AR201455.1 GI:20252343
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE
1 (bases 1 to 3984)
Unclassified.
AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 3 19-MAR-2002;
FEATURES
Location/Qualifiers
1..3984
/organism="unknown"
BASE COUNT 1280 a 847 c 798 g 1059 t
ORIGIN

Query Match 24.1%; Score 760; DB 6; Length 3984;
Best Local Similarity 63.7%; Pred. No. 2.7e-158;
Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

Qy 985 CAACCTGAGCCCTCACTGCCGCCGCCCAACACGGTCAACGGCTGACGGCCGTCAGCGCCGACACATC 1044
Db 1653 CACTCTGCTAAATTCAGCGCTTCTGCTACACAGCTCACTGCAATTACACGACTCCTTAFT 1712
Qy 1045 CTTCAACCTGTGAAGAGCGTGTGCTGCGGCCGAGCCCGACAGACCGCTGGGGCCC 1104
Db 1713 CAACAGCGTGAAGATGCGTGTATTAAACAGAGCTACAGACTGCTTAGCACCT 1772
Qy 1105 TCGACCCCGCTGGGTGGCTGCGCCACAGACACCTGCGACTGAGAGCCTGGAGACGAAG 1164
Db 1773 ACACACCCCTCTTGGATACACAGCCCAATTCAAACTGTTCAACCCAGTCCTTTCTGTAG 1832
Qy 1165 GAGGCAGCGCAGGCCACACCCGCTGGATGTGGACTATGGCGGCTCCGAGCGCAGGTGC 1224
Db 1833 GGAACCGCTTCAATGTGACTGTGATGCGACCTGTGTGTAAGCTCCAACTATCAAGGA 1892
Qy 1225 CCACCGCTTCCGTATCCAAAGCACTGTGCTGCCCGCAAGTCTGAGCAGTACAGCGTG 1284
Db 1893 CCACCAACACCTACCCAAACATCTGTGCGACAAAACCCATCTGTTCTCCATACGAG 1952
Qy 1285 GACCTGGACAGCCTGTGCACCAAGTGTGCAGCAGAGTGTGCGAGGGGGCAGCTGATAGAC 1344
Db 1953 TCAATCAGTAAGCCTAGCAAGAGGATCAGCCAAGCTTGCCCAAGGAAGATCAGAGTGAA 2012
Qy 1345 GGGAGTGACAGGCCACAAAGGTGCGAAGGAGAGACAAGCTGGCAGAGACAAAAGCAG 1404
Db 2013 AAGAGTTA-----TGAAATGTGTAGTGGGGATAAAGAAAGAAACAG 2057
Qy 1405 ATTACAGCCTCCCGGTGCTGTCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
Db 2058 ATTACACTTCACTTACTGTTAGGAAAACAAGAAAGATGAAGCGGAAGGAATCT 2117
Qy 1465 CGCATCAAGAGTTACTCCCGCTTATGCCCTTCAAAATCTTCATGGAGCAACACGTTGAGAT 1524
```

```
Db 2118 CGTATCAAGATTATTTCTCTCAAGCATTTAAATCTTTTATGGAGCAACATCTAGAAAT 2177
Qy 1525 GTCAATCAAAACCTACCAGCAGAGGTTCACCCGAGGCTACAGCTGGAGCAGAAATGGCC 1584
Db 2178 GTACTCAAACTCTCAGCAGCGCTTACATCGTAAAAAACAATTAGAAGATGAATGATG 2237
Qy 1585 AAAGTGGGCTCTGTGAGCGCGAGCAGGAGAGATGAGGAAGATCTCTTACCAGAGGAG 1644
Db 2238 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAA 2297
Qy 1645 TCTAACTACAACCGCTGAAGAGGCCCAAGATGGCAAGTCCATGTTTGTGAAAATCAAG 1704
Db 2298 TCTAATTACATCGCTCTTAAAAGGGCTAAAATGGACAAGTCTATGTTTCTGAAGATAAG 2357
Qy 1705 ACTCTAGGATCGGTGCGCTTTGGGGAAGTGTGCTGCTGTTGAAGCTGACACTCACGCT 1764
Db 2358 ACACCTAGGAATAGGAGCATTTGGTGAAGTCTCTTAGCAAGAAAGATAGATCTAAGGCT 2417
Qy 1765 CTGTACGCCCATGAAGACTCTCAGGAAGAGAGTGTCTTGAACGGGAATCAAGTGGCCAT 1824
Db 2418 TTGTATGCAACAAAACCTCTTGAAGAAGAGATGTTCTTTCGAATCAAGTGCCTCAT 2477
Qy 1825 GTCAAGGCTGAGAGGACATCTGCTGAAGCAGACAAATGAGTGGTGGTCAAACTCTAC 1884
Db 2478 GTTAAAGGCTGAGAGATATCTGCTGAAGCTGACAATGAATGGTAGTCTGCTATAT 2537
Qy 1885 TACTCTTCCAGCAAGACAGACCGCTGTACTTTGTGATGAGTACATACCAGCGGGGAT 1944
Db 2538 TATTCATTCCAAGATAAGGACAAATTTATCTTTGTAATGGACTACATTTCTGGGGTGAT 2597
Qy 1945 ATGATGAGCTGCTGATCAGGATGGAGGTCTTCCCTGAGCAGCTGGCCCGCTTCTACATT 2004
Db 2598 ATGATGAGCTTATTAATAGAAATGGCATCTTTCAGAAAGTCTGGCAGCATTTACATA 2657
Qy 2005 GCAGAGTTGACCTTGCCCATTCGAAAGTGTCCAAAGTGTTCACACGAGTGGTTCACCGGACATC 2064
Db 2658 GCAGAACTTACCTGTGACGTTGCAAGTGTTCATAAAATGGGTTTATTCATAGAGATAFT 2717
Qy 2065 AGCCTGTCAACACATCTACTCATCCACCTGGATGTGCATATTAAGCTGACAGATTTGGCCTC 2124
Db 2718 AAACCTGATAATTTTGTGATGCTGATGCTCATATAATTAATGACTGACTTTGGCCTC 2777
Qy 2125 TGCACCTGGAATCAGGTGGACTCACAATTCGAAGTACTACAGAAAGGGAACACATGAGA 2184
Db 2778 TGCACCTGCTTCAAGTGGACACACGATCTAAGTACTATCAGAGTGGTGAACCATCCAG 2837
Qy 2185 CAGGACAGACTGGAGCCCGGTGACCTCTGGGACGATGTTTCCAACTGTGCTGTGGAGAC 2244
Db 2838 CAAGATAGCATGGATTTCAAGTAATGAATGGGGGATCCCTCAAGCTGTGATGTGGAGAC 2897
Qy 2245 AGGTTAAAGACCTGGAGCAGAGGGCGGAGAGCAGCAGCAGGAGTGGCTGGCAGATTTCT 2304
Db 2898 AGACTGAAGCCCATTAGCGGAGAGCTGCACGCCACAGCAGCGATGTCTAGCAGATCT 2957
Qy 2305 CTGTGCGGACACCAAAATACATCGCTCGGAGGTCTTCTCCGAAAGGGTACAGCAG 2364
Db 2958 TTGGTTGGACTCCCAATTAATTTGCACCTGAAGTGTGCTACGACAGGATACACAG 3017
Qy 2365 CTCTGTGACTGTGGAGCGTGTGATGTTCTTTGAGATGCTGTTGGGAGCGCGCT 2424
Db 3018 TTGTGTGATTTGGTGGAGTGTGGTGTATTCTTTTGAATGTTGTGGGACAACCTCTCT 3077
Qy 2425 TTCTTTGGCCCCCCCCACACAGCAGCGAGCTGAAGGTGATCAACTGGGAGAGCAGCTG 2484
Db 3078 TTCTTTGGCAAAACACCAATTAGAAAACAATGAAGGTATCAACTGGCAAACTCTCT 3137
Qy 2485 CATATCCCTTACCGAGTGAAGCTCAGCGCTGAGCGCTGAGCGCCGAGACCTCATCAGAGCTGTC 2544
Db 3138 CACATTCCACCAACAGCTAAACTCAGTCTGAGCTCTGATCTTATTATTAACTTTGC 3197
Qy 2545 TCGCGGCTGACTGCGGCTGGGCGAGGATGGGCGAGATGACCTCAAGSCACACCGCTTC 2604
```

```

Db 3198 CGAGGACCCGAGATCGCTTAGGCAAGATCGTGTGATGAATAAAGAGTCATCCATTT 3257
Qy 2605 TTCAACACCATCGACTTTTCCGTCGATCCGAAGAGCGGTGCACCCCTACGTCGCCACC 2664
Db 3258 TTTAAACAATTGACTTCTCCAGTACGCTGAGACAGAGTCTGCTTCATACATTCCTAAA 3317
Qy 2665 ATCAGCACCCATCGGACACCTTCAATTTTACCGCGTGGATGAAGAAGCCCTGGCAC 2724
Db 3318 ATCAGCACCCATCGGACACCTTCAATTTTACCGCGTGGATGAAGAAGCCCTGGCAC 3377
Qy 2725 GAGGCCAGC---GGAGAGAGCGCCAAAGCCCTGGGACAGCGTGGCTCCGCCAGCAGCAAG 2781
Db 3378 GATGATACGAGGAGAAATGTAATGACACTCTCAATGATGATGATAAATGGAAG 3437
Qy 2782 CATCAGACGACGCTTCTATGATTCACCTTCGCGAGGTTCTTCGATGACAAAGCGCTAT 2841
Db 3438 CATCCTGAACATCGATCTATGAATTTACCTTCGGAAGTTTTCGATGACAAATGGCTAC 3497
Qy 2842 CCCTTCGGTCCCGAAGCC 2861
Db 3498 CCATATAATTATCCGAAGCC 3517

RESULT 10
AF164041
LOCUS Homo sapiens WARTS protein kinase (WARTS) mRNA linear PRI 06-DEC-1999
DEFINITION AF164041
ACCESSION AF164041
VERSION AF164041.1 GI:5738135
KEYWORDS
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 4241)
AUTHORS Kishimoto,T., Niwa,S.-i., Nagamine,Y., Nishiyama,Y. and Saya,H.
TITLE WARTS protein, polynucleotide encoding the same, antisense
polynucleotide thereof, and antibody recognizing the protein
JOURNAL Patent: Japan (PCT/JP98/03739) 24-AUG-1998;
Sumitomo Electric Industries, Ltd.; 1 Taya-cho, Sakae-Ku,
Yokohama-shi; Kanagawa;
Japan;
REFERENCE 2 (bases 1 to 4241)
AUTHORS Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Yamamoto,H., Hiraoka,T., Kitamura,N. and
Saya,H.
TITLE A human homolog of Drosophila warts tumor suppressor, h-warts,
localized to mitotic apparatus and specifically phosphorylated
during mitosis
JOURNAL FEBS Lett. 459 (2), 159-165 (1999)
MEDLINE 99447636
PUBMED 10518011
REFERENCE 3 (bases 1 to 4241)
AUTHORS Nishiyama,Y., Hirota,T., Morisaki,T., Hara,T., Marumoto,T.,
Iida,S., Makino,K., Nakamura,H., Koga,H. and Saya,H.
TITLE Direct Submission
JOURNAL Submitted (30-JUN-1999) Tumor Genetics and Biology, Kumamoto
University School of Medicine, 2-2-1 Honjo, Kumamoto 860-0811,
Japan

FEATURES
Source Location/Qualifiers
1. .4241
/organism="Homo sapiens"
/db_xref="taxon:9606"
gene 1. .4241
/genes="WARTS"
CDS 33..3425
/genes="WARTS"
/note="similar to Drosophila warts tumor suppressor;
localized on mitotic apparatus, specifically
phosphorylated at mitotic checkpoint"
/codon_start=1
/product="WARTS protein kinase"
/protein_id="AAD50272.1"

```

```

/db_xref="GI:5738136"
/translation="MKRSEKPEGYRQMRPKTFPPASNTVTVSSRMLQIRESLRNLSP
SDAAKHNNKSMKSTEDPROVRNPPKFGTHKALQELRNSLLPFANETNSRSTSEVN
POMLODLOAAGFEDDMVIOALKTNNRSIEAAEFISKSMYQDPDROMAAAAARPIN
ASKMPPGNVOQSVNRKOSKESLVPORHGPPLGESVAVHSESPNSQTDVGPPLSGS
GISAFVQAHNSGNQRVNPDPQVRSVTPPPPPGQTPPPRPTTPPPSPWEPNSQTKR
YSKMEYVTSRISPSVPPGAWOEGYPPPLNTSPMNPNGORGISVVPVQSPILMOS
SSKNEFPSPRGMNGOTDFMIHQNVVAGTVNPPPPYPLTAAAGOSPALOTG
GSAAPSYTNGSIPQSMVPPNRNHNMLYNI5VPGLOTNPQSSAPASPSGSGHE
IPTWOPNIPIVRSFNPLNASHANSOPSATVTITPAIQQPVKSMVLPKPEL
OTALATHPSNIPQPIQTVQSPPEBCTASNTVMPVPAEAPNYPQPPYPKHLHQ
NPSPYVESISKSEKEDQSPPKDESEKSYENVDGDKKKQITTSPIVVRNKKDE
ERESRQISVSPQAFKFMHQHVENLKHQHRKKOLENEMVRMVKYALYATLIRKD
KMLQKESNYIRLKRMDKSMFKIKTIGIFAGFEYCLARKYDTKALYATLIRKD
VLLRNOVHVKAEARDILAEADNEMVRLYY5FQDKNLYFVMDYIPGDMMSLIRMG
IFFESLARFYIAELTCAVESVHKMGFHRDIKPDNILDIDRDKIKITDFGLCTGRWT
HDSKYVSGDHPQDSMDFSNEMGDPSSCRDRLKPLERAAARHORCLAHSLGTPP
NYTAPEVLLRTGYTQCDWWSGVILFEMLVGQPPFLAQTPLETOMKVINWQISLHIP
POAKLSPEASDLIIKICRGPDRLGNKGADEIKAHFFKTFIDFSSDLRQOSASYPKI
THPTDTSNFDVPDKLWSDNNEENVDTLNGYKNGKHPHEFYETFRFRFDDNG
YPYNYPKPIEYIINSQSSQSDDEDDONTSETKNRDLVYV"
misc_feature 2145..3059
/genes="WARTS"
BASE COUNT 1384 a 884 c 817 g 1156 t
/notes="Region: serine/threonine kinase domain"
Query Match 24.1%; Score 760; DB 9; Length 4241;
Best Local Similarity 63.7%; Pred. No. 2.8e-158;
Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;
Qy 985 CAACCTGAGCCCTCACATGCGCCGCCCAACACGCTGACGCGCGCGCACACATC 1044
Db 1455 CACTCTGCTAATTCAGCCCTTCTGCTACACAGCTACTGCAATTACACAGCTCTATT 1514
Qy 1045 CTTCCACCTGTGAAGAGCGTGGTGTGCGGCGCCGAGCGCCGAGCGTGGGCGCC 1104
Db 1515 CAACAGCCTGTGAAAGTATGCGTGTATTAAACACGAGAGCTACAGACTGCTTTAGCACCT 1574
Qy 1105 TCGCACCCGCGCTGGTGGTGGCGCCACAGCCTGCCACTGAGAGCGCTGGAGCAGG 1164
Db 1575 ACACACCTCTCTGGATACACAGCCAAATTCAACTGTTCACCCAGTCTCTTCTCTGAG 1634
Qy 1165 GAGGCGAGCGAGCCACACACCGCTGGATGTGGAGTATGCGGCTCCGAGCGCAGGTGC 1224
Db 1635 GGAACCCCTTCAAATGTGACTGTGATCCACTCTTCTGTAAGCTCCAAACTATCAAGGA 1694
Qy 1225 CCACCGCTTCGGTATCAAAGCACTTCTGCTGCCAGTAGTCTGAGCAGTACAGCGTG 1284
Db 1695 CCACCAACCCCTACCCAAACATCTCTGTCACCAAAACCCATCTCTTCTCTCCATACGAG 1754
Qy 1285 GACCTGCACAGCCTGTGCACAGTGTGCAGAGTCTGGAGGGGCGCTGATCTAGAC 1344
Db 1755 TCATCATAGTAACTAGCAAGAGGATCAAGCAAGCTTGCCCAAGGAAGATGAGAGTGAA 1814
Qy 1345 GGGAGTGACAGAGCCACAAAGGTGCGAAGGGAGACAAAGCTGGCAGAGACAAAGACAG 1404
Db 1815 AAGAGTTA-----TGAAATGTTGATAGTGGGATAAAGAAAGAAACAG 1859
Qy 1405 ATTCAGACCTCCCGGTGCTGTCCGCAAGAAATAGCAGATGAAGAGAGAGAGAGTCT 1464
Db 1860 ATTACAACTTCACCTATTACTGTAGGAAAACAAAGAAAGATGAAGAGCGAAGGAATCT 1919
Qy 1465 CGCATCAAGAGTTACTCCCTTATGCTTCAAAATTTCTTCATGGAGCAACAGCTGGAGAA 1524
Db 1920 CGTATCAAGATTATCTCTCTCAAGCAATTTAAATCTTTATGGAGCAACATGTAGAAAT 1979
Qy 1525 GTCATCAAAACCTACCAAGAGGTCAGCGGAGGCTACAGCTGGAGCAAGAAATGCC 1584
Db 1980 GTACTCAAACTCATCAGAGCGCTACATCGTAAAAAACAATTAGAGAAATGAATGATG 2039
Qy 1585 AAGCTGGGCTCTGTGAGGCGGAGCAGCAGATGAGGAAGATCTCTTACCAGAGGAG 1644

```


QY 2842 CCCTTCGGTGGCCGAAGCC 2861
Db 3816 CCATATAATTATCCGAAGCC 3835

RESULT 12
E24613 2442 bp DNA linear PAT 18-JUN-2001
DEFINITION warts protein, polynucleotide encoding the protein, antisense
polynucleotide thereof and antibody recognizing the protein.
ACCESSION E24613
VERSION E24613.1 GI:13024646
KEYWORDS JP 1999089580-A/1.
SOURCE unidentified.
ORGANISM unidentified.
REFERENCE 1 (bases 1 to 2442)
AUTHORS Toshihiko K., Shinichiro, T., Kyoko, N., Yasuyuki, N. and Hideyuki, S.
TITLE warts protein, polynucleotide encoding the protein, antisense
polynucleotide thereof and antibody recognizing the protein
JOURNAL Patent: JP 1999089580-A 1 06-APR-1999;
SUMITOMO ELECTRIC IND LTD
COMMENT OS Unidentified
PN JP 1999089580-A/1
PD 06-APR-1999
PF 24-SEP-1997 JP 1997258689
PR
PI TOSHIHIKO KISHIMOTO, SHINICHIRO TAMBA, KYOKO NAGAMINE, PI
YASUYUKI NISHIYAMA,
PI HIDEYUKI SAYA,
PC C12N15/09, C07K14/435, C07K16/18, C1201/68, G01N33/53, PC
G01N33/532, C12P21/02,
PC (C12N15/09, C12R1:91), (C12P21/02, C12R1:19), C12N15/00,
(C12N15/00, PC C12R1:91)
CC Strandedness: Double;
FT source 1. .2442
FT key Location/Qualifiers
FT source 1. .2442
FT Location/Qualifiers
1. .2442
/organism='Unidentified'.
/db_xref='taxon:32644'

BASE COUNT 777 a 528 c 491 g 646 t

Query Match 23.7%; Score 747.8; DB 6; Length 2442;
Best Local Similarity 63.5%; Pred. No. 1.4e-155;
Matches 1184; Conservative 0; Mismatches 562; Indels 19; Gaps 2;

QY 985 CAACCTGAGCCCTCACTGCGCCGCCACACGTCACGCGTGCAGCGCCGCGCACACATC 1044
Db 592 CACTCTGCTAATTCAGCCCTTCTGCTACACAGTCACTGCAATACACGCTCCTATT 651

QY 1045 CTTCACTGCTGAAGAGCGTGTGTGTCGCGCCGCCAGACGCCCTGGGGCC 1104
Db 652 CAACAGCCTGTCAAAGTATGCTGTATTAACACAGACTACAGACTGCTTTAGCACCT 711

QY 1105 TGCACCCCGCTGGTGGCTGCGCCGCCACAGCACTGCCACTGAGAGCGCTGGAGACGA 1164
Db 712 ACACACCCCTTCTGGATACACAGCCCAATTCAAACTGTTCACCCAGTCTTTCTGTAG 771

QY 1165 GAGGCGAGCGGAGCCACACCCGCTGGATGTGACTATGGCGCTCCGAGCGAGTGC 1224
Db 772 GGAACCGCTTCAATGTGACTGTGATGCCACTGTGCTGGAAGCTCCAACTATCAAGA 831

QY 1225 CCACCGCTTCCCTATCAAAGCACTGTGCTGCGCCAGTAACTGTGAGCAGTACAGCGTG 1284
Db 832 CCACACCACTTCCCAAAACATCTGTGACCAAAACCCATCTGTTCTCCATACAG 891

QY 1285 GACCTGGACAGCCTGTGCACCACTGTGCAGCAGAGTCTGCGAGGGGCGCACTAGAC 1344
Db 892 TCAATCAGTAAGCCTAGCAAGAGGATCAGCCCAAGCTTGCCCAAGGAAGATGAGAGTGA 951

QY 1345 GGGAGTGCACAGAGCCACAAAGTGGAGGAGACAAAGCTGGCAGACACAAAAGCAG 1404
Db 952 AAGAGTTA-----TGAAATGTTAGTGGGATAAAGAAAGAACAG 996

QY 1405 ATTCAGACCTCCCGTGCCTGTCCGCAAGATAGCAGAGATGAAGAGAGAGAGTCT 1464
Db 997 ATTCAACCTTCACTTACTCTTAGGAAAAACAAGAAAGATGAAGAGCGAAGGAACT 1056

QY 1465 CCATCAAGAGTTACTCCCTTATGCTTCAATTTCTTCATCGAGCAACACCTGGAGAT 1524
Db 1057 CGTATTCAAAGTTATTTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATAGAAA 1116

QY 1525 GTCAACAAACCTACACAGAGAGTGCAGCGAGGCTACAGCTGCAGCAGAAATGGCC 1584
Db 1117 GTACTCAATCTCAGCAGCGCTCTACATCTGTAACAAACAATTAGAGATGAATGATG 1176

QY 1585 AAAGCTGGGCTCTGTGAGCCGAGCAGAGATGAGGAAGATCTCTTACAGAGAGG 1644
Db 1177 CGGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCCAAGAA 1236

QY 1645 TCTAATCAACCGCTGAAGAGGCGCAAGATGGACAGTCCATGTTTGTGAAATCAAG 1704
Db 1237 TCTAATTACATCCGCTTTAAAAAGGCTAAATGGACAAGTCTATGTTTGTGAAGATAAG 1296

QY 1705 ACTCTAGGATCGGTGCTTGGGGAAGTGCCTCGCTTGTAAAGCTGCACACTCACGCT 1764
Db 1297 ACACTAGGAATAGGAGCATTTGTTGAAGTCTGTCTAGCAAGAAAGATAGTACTTAGGCT 1356

QY 1765 CTGTACGCCATGAAGACTCTCAGGAAGAGAGATGCTCTGAACCGGAATCAAGTGGCCAT 1824
Db 1357 TTGTATGCAACAAAACCTCTTCCGAAAGAAAGATGTTCTTTCGAAATCAAGTGCCTCAT 1416

QY 1825 GTCAAGCTGAGAGGAGCATCTCTGCTGAAGCAGACAAATGAGTGGTGTCAACTCTAC 1884
Db 1417 GTTAAAGCTGAGAGAGATATCTCTGCTGAAGCTGCAATGAATGGGTAGTCTCTATAT 1476

QY 1885 TACTCTTCCAGGACAAAGGACAGCTGTATTTGTGTGAGTACTACATACACGCGGGAT 1944
Db 1477 TATTCATTCAGAGTAAGGACAAATTTACTTTGTAAGGACTACATCTCTGGGGTGTAT 1536

QY 1945 ATGATGAGCCTGCTGATCAGGATGGAGTCTTCCCTGAGCAGCTGGCCGCTTCTACAT 2004
Db 1537 ATGATGAGCCTTATTAATAGAAATGGCATCTTTCAGAAAGTCTGCGCAGATTTCTACATA 1596

QY 2005 GCAGAGTTGACCTGGCCATTCGAAAGTGTCCACAGATGGGCTTTATCCACCGGACATC 2064
Db 1597 GCAGAACTTACCTGTGCGAGTTGAAAGTGTTCATAAATGGGTTTTTATTTCATAGATAT 1656

QY 2065 AAGCCTGACACATCTATCTGACCTGGATGCTATATTAAGCTGACAGATTTTGGCCCTC 2124
Db 1657 AAACCTGATATATTTTGTGATGCTGATGCTATATTAATTTGACTGACTTTGGCCCTC 1716

QY 2125 TGCATGGATTCAGGTGGACTCACAATTCGAAAGTACTACAGAAAGGGAACACATAGAGA 2184
Db 1717 TGCATGGCTTCAGATGGACACACGATCTAAGTACTATCAGAGTGGTGACCATCCACGG 1776

QY 2185 CAGGACAGATGGAGCCCGGTGACCTCTGGGAGGATGTTTCCAACTGTCGCTGTGGAGAC 2244
Db 1777 CAAGATAGCATGGATTTTCAATTAATGGGGGGATCCCTCAAGCTGTCGATGTGGAGAC 1836

QY 2245 AGGTTAAAGACCTTGAGCAGAGGCGGAGAGAGCAGCAGAGTGGCTGGCAGCATTTCT 2304
Db 1837 AGACTGAAGCCATTAAGCGGAGAGCTGACCCAGCAGCAGGATGCTACACATTTCT 1896

QY 2305 CTTGTGGGACACCAAAATTAATCTCGCGAGGTGCTTCTCCGAAAGGGGTACACGAG 2364
Db 1897 TTGGTTGGGACTCCCAATTAATTCACCTGAAGTGTGCTACGAAACAGGATACACAG 1956

QY 2365 CTCCTGTGACTGGTGGAGCGTGGTGTGATCTCTTTTGTGAGATGCTGGTGGGCGCCGCT 2424
Db 1957 TTGTGTGATGGTGGAGTGTGGTGTATTTCTTTTGTGAAATTTGTTGGTGGGACAACTCT 2016

Moser, M., Neal, D., Newton, J., Newton, N., Nguyen, A., Nguyen, N.,
Nguyen, N., Nickerson, E., Nwokenko, S., Ogih, M., Okuwona, G.,
Oragunye, N., Oviedo, R., Pace, A., Payton, B., Peery, J., Perez, L.,
Peters, L., Pickens, R., Primus, E., Pu, L.L., Quiles, M., Ren, Y.,
Rives, M., Rojas, A., Rojibokan, I., Rolfe, M., Ruiz, S., Savery, G.,
Scherrer, S., Scott, G., Shen, H., Shoostari, N., Sisson, I.,
Sodergren, F., Sotaike, T., Sparks, A., Stanley, H., Stone, H.,
Sutton, A., Svatek, A., Tabor, P., Tamerisa, A., Tamerisa, K., Tang, H.,
Tansey, J., Taylor, C., Taylor, T., Telford, B., Thomas, N., Thomas, S.,
Usmani, K., Vasquez, L., Vera, V., Villalón, D., Vinson, R., Wang, Q.,
Williams, S., Ward-Moore, S., Warren, R., Washington, C., Watlington, S.,
Williams, G., Williamson, A., Wleczyk, R., Wooden, S., Worley, K.,
Wu, C., Wu, Y., Wu, Y.F., Zhou, J., Zorrilla, S., Nelson, D.,
Weinstock, G. and Gibbs, R.
Direct Submission
Unpublished
2 (bases 1 to 57393)
Worley, K.C.
Direct Submission
Submitted (15-SEP-2001) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
3 (bases 1 to 57393)
Worley, K.C.
Direct Submission
Submitted (23-JUL-2002) Human Genome Sequencing Center, Department
of Molecular and Human Genetics, Baylor College of Medicine, One
Baylor Plaza, Houston, TX 77030, USA
On Jul 17, 2002 this sequence version replaced gi:17941559.

----- Genome Center -----
Center: Baylor College of Medicine
Center code: BCM
Web site: <http://www.hgsc.bcm.tmc.edu/>
Contact: hgsc-help@bcm.tmc.edu
Project Information
Center project name: GBJE
Center clone name: CH230-401
Summary Statistics
Chemistry: Dye-terminator Big Dye: 100% of reads
Assembly program: Phrap: version 0.990329
Consensus quality: 38843 bases at least Q40
Consensus quality: 42803 bases at least Q30
Consensus quality: 45893 bases at least Q20

* NOTE: Estimated insert size may differ from sequence length
(see http://www.hgsc.bcm.tmc.edu/docs/genbank_draft_data.html)
* NOTE: This is a 'working draft' sequence. It currently
* consists of 33 contigs. The true order of the pieces
* is not known and their order in this sequence record is
* arbitrary. Gaps between the contigs are represented as
* runs of N, but the exact sizes of the gaps are unknown.
* This record will be updated with the finished sequence
* as soon as it is available and the accession number will
* be preserved.

1	1327: contig of 1327 bp in length
1328	1427: gap of unknown length
1428	2480: contig of 1053 bp in length
2481	2580: gap of unknown length
2581	3874: contig of 1294 bp in length
3875	3974: gap of unknown length
3975	6236: contig of 2262 bp in length
6237	6336: gap of unknown length
6337	7588: contig of 1252 bp in length
7589	7688: gap of unknown length
7689	9085: contig of 1397 bp in length
9086	9185: gap of unknown length
9186	10268: contig of 1083 bp in length
10269	10368: gap of unknown length
10369	11925: contig of 1557 bp in length
11926	12025: gap of unknown length
12026	13508: contig of 1483 bp in length
13509	13608: gap of unknown length
13609	15071: contig of 1463 bp in length

1

Fri Jan 17 11:17:02 2003

QY 2514 TGAGGCCGAGACCTCATCAGCAAGCTGTCTCGCGGCTGACTGCCGCTGGGAGGA 2573
 Db 4088 CGAGGCTACGGACTTGTATAGGAGGCTCTGTGGTGGCTGACAAGCGGCTGGCA--A 4144
 QY 2574 TGGGGCAGATGACCTCAAGGCACACCGCTTCTCAACACCATCGACTTTCCCGTGACAT 2633
 Db 4145 GAGCGTGGACGAGGTCAAGAGCCACGACTTCTTCAAGGCGATCGACTTTGC---GGACAT 4201
 QY 2634 CCGAAGCAGGCTGCACCTACCTCCACCATCAGCCACCTCCCAATTT 2693
 Db 4202 GCGAAGCAGAAAGCGCCCTACATCCGGAAATCAAGACCCCAACGACATCCCACTT 4261
 QY 2694 TGACCCGCTGATGAGAAAGCCCTTGGCACAGGCGCAGGAGAGCGGCCAAGGCCTG 2753
 Db 4262 TGATCCGCTGATCCGGAAGGCTGGCTCGAATGACTCCACCATGAGCAGCGCGATGA 4321
 QY 2754 GGACACGCTGGCCTCCCGCAGCAGCAGCATCCAGAGCGGCTTCTATGAGTTACCTT 2813
 Db 4322 TGTGACACAGAAATGACCGCA-----CTTCCAGCGCTTTTCAATTTACCTT 4369
 QY 2814 CCGCAGGTTCTTCGATGACAA 2834
 Db 4370.CCGTCGCTTCTTCGACGACAA 4390

Search completed: January 16, 2003, 18:31:45
 Job time : 5402.57 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run On: January 16, 2003, 09:50:07 ; Search time 402.423 Seconds
(without alignments)
17655.667 Million cell updates/sec

Title: us-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagagccaccgcgaagt.....aagagcactattttggggg 3155

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002.*
1: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
2: /SID52/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SID52/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SID52/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SID52/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SID52/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SID52/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SID52/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SID52/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SID52/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SID52/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SID52/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SID52/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SID52/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SID52/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SID52/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SID52/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SID52/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SID52/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SID52/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SID52/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SID52/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SID52/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3155	100.0	3155	17 AAT42120	M-lats2 gene encod
2	3155	100.0	3155	21 AAZ51507	Mouse Lats2 (large
3	1749.6	55.5	5276	20 AAX87397	Human WART2 cDNA.
4	1749.6	55.5	5486	21 AAX59129	DNA encoding a tum
5	1748	55.4	5486	21 AAX59130	DNA encoding a tum
6	1717.2	54.4	3533	22 AAT17165	Human cancer-inhib
7	1293	41.0	1961	21 AAZ61160	CDNA SS1771A encod
8	1148.4	36.4	1498	21 AAZ61158	CDNA SS1771 encodi
9	1118.6	35.5	1912	23 ABK43457	DNA encoding novel

10	1035.2	32.8	2043	22 AAD03989	Human protein tyro
11	812.6	25.8	3213	21 AAZ51506	Mouse Lats (large
12	809.4	25.7	3213	17 AAT42119	M-lats gene encodi
13	760	24.1	3984	17 AAT42118	H-lats gene encodi
14	760	24.1	3984	21 AAZ51505	Human Lats (large
15	760	24.1	7382	20 AAX87396	Human WART1 cDNA.
16	758.8	24.1	1501	22 AAF75341	Human TGF-beta rec
17	747.8	23.7	2442	20 AAX32981	Human warts protei
18	676.6	21.4	1374	20 AAX32982	Human warts gene f
19	567.8	18.0	3319	23 ABL03169	Drosophila melanog
20	564.6	17.9	5720	17 AAT42117	Lats gene encoding
21	564.6	17.9	5720	21 AAZ51508	Drosophila melanog
22	551.8	17.5	1357	22 ABA08740	Human large tumour
23	546.6	17.3	676	22 AAS27179	CDNA encoding nove
24	546.6	17.3	676	23 ABK43776	DNA encoding novel
25	333	10.6	638	21 AAZ79942	Human colon cancer
26	258.8	8.2	2160	22 AAI66703	A. gossypii AG007
27	248.2	7.9	11187	23 ABL03168	Drosophila melanog
28	241.8	7.7	2255	23 ABL29695	Drosophila melanog
29	237	7.5	4983	22 AAF44629	Novel protein kina
30	235	7.4	1710	21 AAC43403	Arabidopsis thalia
31	235	7.4	2003	21 AAC49841	Arabidopsis thalia
32	233.8	7.4	1935	20 AAX06834	Disease associated
33	231.6	7.3	1894	21 AAC39567	Arabidopsis thalia
34	230	7.3	1922	21 AAC49426	Arabidopsis thalia
35	225.2	7.1	3018	17 AAT31452	Human Ndr serine/t
36	225.2	7.1	3018	24 ABK84773	Human cDNA differe
37	225.2	7.1	3800	22 ABA08668	Human NDR homologo
38	222.6	7.1	2101	17 AAT31451	Drosophila Ndr ser
39	220.2	7.0	2001	21 AAC36174	Arabidopsis thalia
40	214.4	6.8	678	22 AAZ79982	Human colon cancer
41	192.2	6.1	734	22 AAS27161	CDNA encoding nove
42	192.2	6.1	734	22 ABK56728	Human immune/haema
43	192.2	6.1	734	23 ABK43709	DNA encoding novel
44	192.2	6.1	734	23 ABK43985	DNA encoding novel
45	166.4	5.3	1788	19 AAV71036	mPKAC-green floure

ALIGNMENTS

RESULT 1
AAT42120
ID AAT42120 standard; cDNA: 3155 BP.

XX AAT42120;

AC AAT42120;

XX 31-JAN-1997 (first entry)

DT M-lats2 gene encoding large tumour suppressor.

DE Mouse; m-lats2 gene; large tumour suppressor; fetal brain;
protein-serine/threonine-kinase; cell proliferation; antisense;
dominant-negative; cancer; degenerative disorder; trauma;
growth deficiency; therapy; antitumour; vulnery; diagnostic;
transgenic plant; transgenic animal; growth; senescence; ds.

KW Mus musculus.

OS Mus musculus.

XX Key Location/Qualifiers

FT CDS 1..2943

FT FT /*tag= a

FT FT /product= m-lats2 protein

XX W09630402-A1.

PN Mouse Lats2 (large

XX Human WART2 cDNA.

XX 03-OCT-1996.

XX 26-MAR-1996; 96WO-US04101.

XX 27-MAR-1995; 95US-0411111.

XX (UYVA) UNIV YALE.

XX Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX WPI: 1996-455275/45.
 DR P-PSDB; AAW06287.
 XX
 PT New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX
 PS Claim 8; Page 133-137; 215pp; English.
 XX
 CC This sequence encodes a mouse large tumour suppressor m-lats2
 CC protein, and has been isolated from a newborn mouse brain phase
 CC lambda-ZAP cDNA library using a 2.2-kb DNA probe from the Drosophila
 CC lats gene (AAT42117). A homologous mouse sequence has also been
 CC isolated (m-lats, AAT42119). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 SQ Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;

Query Match 100.0%; Score 3155; DB 17; Length 3155;
 Best Local Similarity 100.0%; Pred. No. 0;
 Matches 3155; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 ATGAGAGCCACCCGAGTTTGGACCTTATCAAAAAGCTCTCAGGAAATCCGATATTC 60
 DB 1 ATGAGAGCCACCCGAGTTTGGACCTTATCAAAAAGCTCTCAGGAAATCCGATATTC 60
 QY 61 CTCCTGCTTTTCCACAGGAGTTCAGGACCTTCGCGACCTCGCAGAGTGACCGGCAGATG 120
 DB 61 CTCCTGCTTTTCCACAGGAGTTCAGGACCTTCGCGACCTCGCAGAGTGACCGGCAGATG 120
 QY 121 CTTGAGGAGTTGGTGAATCGGCGATGTGACAGAGAGATGGCTGGCAGAGCGCTCACGACG 180
 DB 121 CTTGAGGAGTTGGTGAATCGGCGATGTGACAGAGAGATGGCTGGCAGAGCGCTCACGACG 180
 QY 181 ACGGCAGTAGGAGTATCGAGTGCCTTGGAGTACATAGTACATAGTGGCTACCTGGAC 240
 DB 181 ACGGCAGTAGGAGTATCGAGTGCCTTGGAGTACATAGTACATAGTGGCTACCTGGAC 240
 QY 241 CCCAGGAATGAGCAGATGTGTGCGAGTTCATCAAGCAGACCTCCCGCAGAAAGGCGCTGGCG 300
 DB 241 CCCAGGAATGAGCAGATGTGTGCGAGTTCATCAAGCAGACCTCCCGCAGAAAGGCGCTGGCG 300
 QY 301 TCCACCCGCTGACTCGCGCGCCAGTTTCAGGGCAGACGGGAAAGCACTCCCATCTTAC 360
 DB 301 TCCACCCGCTGACTCGCGCGCCAGTTTCAGGGCAGACGGGAAAGCACTCCCATCTTAC 360
 QY 361 CACGAGTGGTGTGCAACTACGAGGGCCCGCGCGACTGGAGGAGATGGCGCGGAA 420
 DB 361 CACGAGTGGTGTGCAACTACGAGGGCCCGCGCGACTGGAGGAGATGGCGCGGAA 420
 QY 421 TATTTAGACTTTCTTCCCTGGAGCGGAGCGGACCCACCGCTGCCAGGCTCACACG 480
 DB 421 TATTTAGACTTTCTTCCCTGGAGCGGAGCGGACCCACCGCTGCCAGGCTCACACG 480
 QY 481 CATCTCCCAAGAGGTACAGCAGACAGTAGAGCAAGTGGCGACTTTCCGGGACACAC 540
 DB 481 CATCTCCCAAGAGGTACAGCAGACAGTAGAGCAAGTGGCGACTTTCCGGGACACAC 540
 QY 541 TATGTCGTCGTCATCTACTATCGGAGCAGTCTGGGTATGGGTGTCAGCGAGTTCCTCC 600
 DB 541 TATGTCGTCGTCATCTACTATCGGAGCAGTCTGGGTATGGGTGTCAGCGAGTTCCTCC 600
 QY 601 TTCCAGAACAGACGCCACCATGATCTTATCCAGCATGCCCCAAGGCCCGAGGTGGCCCT 660

DB 601 TTCCAGAACAGACGCCACCATGATGCTTATCCAGCATGGCCAAAGGCCGAGGTGGCCCT 660
 QY 661 CCGCCGAGCTTACCTTTCTGTCGCCATGCTGGCTGTACACTGCTCGCACCAAGCCG 720
 DB 661 CCGCCGAGCTTACCTTTCTGTCGCCATGCTGGCTGTACACTGCTCGCACCAAGCCG 720
 QY 721 CGGCTACCCACCTGGGGCCACCCATATACATGTGTGGGACCCCGGGTCCACGTTT 780
 DB 721 CGGCTACCCACCTGGGGCCACCCATATACATGTGTGGGACCCCGGGTCCACGTTT 780
 QY 781 ACTGCGAAAGCTCTGCACAGGCTGTGCTGSCACCGTCCAGGACAGCCCTCAATGCTGAC 840
 DB 781 ACTGCGAAAGCTCTGCACAGGCTGTGCTGSCACCGTCCAGGACAGCCCTCAATGCTGAC 840
 QY 841 TTGTACGAGCTGGGCTCCACGGTGGCTGTGCTGSCAGCTCCACTGGCAGCCGCGACTCG 900
 DB 841 TTGTACGAGCTGGGCTCCACGGTGGCTGTGCTGSCAGCTCCACTGGCAGCCGCGACTCG 900
 QY 901 CTGACAGAGCAGGCTTAGAAGCTTCGCGCGGATGCTGCTTTTGGGGTGGCGCCAGC 960
 DB 901 CTGACAGAGCAGGCTTAGAAGCTTCGCGCGGATGCTGCTTTTGGGGTGGCGCCAGC 960
 QY 961 AGGACCAACTCTTCAAAACCCACAACTGAGCCCTCACTGCCGCCGCCCAACAGCGTC 1020
 DB 961 AGGACCAACTCTTCAAAACCCACAACTGAGCCCTCACTGCCGCCGCCCAACAGCGTC 1020
 QY 1021 ACCGCGTGTGAGCGCGCACATCTTACCCCTGTGAAGAGCGTGTGCTGCGGCC 1080
 DB 1021 ACCGCGTGTGAGCGCGCACATCTTACCCCTGTGAAGAGCGTGTGCTGCGGCC 1080
 QY 1081 GAGCCCGCAGACGCTGGGGCCCTCGCACCCGCTGGGTGGCTGGCGCCACAGCACCT 1140
 DB 1081 GAGCCCGCAGACGCTGGGGCCCTCGCACCCGCTGGGTGGCTGGCGCCACAGCACCT 1140
 QY 1141 GCACCTGTAGAGCCTTGAGAGCGAGGAGGCGAGGCGCCACACCGCTGGATGTGGAC 1200
 DB 1141 GCACCTGTAGAGCCTTGAGAGCGAGGAGGCGAGGCGCCACACCGCTGGATGTGGAC 1200
 QY 1201 TATGGCGCTTCGAGCGCAGGTGCCACCGCTCCGCTATCAAGACACTTGTGTGTC 1260
 DB 1201 TATGGCGCTTCGAGCGCAGGTGCCACCGCTCCGCTATCAAGACACTTGTGTGTC 1260
 QY 1261 AGTAAGTCTGAGCAGTACAGCGTGGACCTGTGACAGCTGTGACAGAGTGTGTCAGAGT 1320
 DB 1261 AGTAAGTCTGAGCAGTACAGCGTGGACCTGTGACAGCTGTGACAGAGTGTGTCAGAGT 1320
 QY 1321 CTGCGAGGGGCTATGATCTAGACGGAGTGACAGAGCCACAAAGGTCGGAAGGAGAC 1380
 DB 1321 CTGCGAGGGGCTATGATCTAGACGGAGTGACAGAGCCACAAAGGTCGGAAGGAGAC 1380
 QY 1381 AAAGCTGGCAGAGACAAAGAGCAGATTTCAGACCTCCCGGTGCTGCGCAAGATAGC 1440
 DB 1381 AAAGCTGGCAGAGACAAAGAGCAGATTTCAGACCTCCCGGTGCTGCGCAAGATAGC 1440
 QY 1441 AGAGATGAAGAGAGAGAGTCTGCGATCAAGAGTACTCCCTTATGCTTCAATTC 1500
 DB 1441 AGAGATGAAGAGAGAGAGTCTGCGATCAAGAGTACTCCCTTATGCTTCAATTC 1500
 QY 1501 TTTCATGGAGCACACGCTGGAGATGTATCAAAACCTTACAGCAGAGGTCAGCGGAGG 1560
 DB 1501 TTTCATGGAGCACACGCTGGAGATGTATCAAAACCTTACAGCAGAGGTCAGCGGAGG 1560
 QY 1561 CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGAGGCGGACAGGAGGATG 1620
 DB 1561 CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGAGGCGGACAGGAGGATG 1620
 QY 1621 AGGAAGATCTCTTACCAGAGAGAGTCTTAACACACCGCTGAAGAGGCGCAAGATGAC 1680
 DB 1621 AGGAAGATCTCTTACCAGAGAGAGTCTTAACACACCGCTGAAGAGGCGCAAGATGAC 1680
 QY 1681 AAGTCCATGTTTGTGAAATCAAGACTCTAGGATCGTGGCTTTGGGGAAGTGTGCTC 1740

Db 1681 AAGTCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGGAAAGTGCCTC 1740
QY 1741 GCTTGTAAAGTGGACACTCAGCCTCTGTACGCCATGAGACTCTCAGGAAGAAGATGTC 1800
Db 1741 GCTTGTAAAGTGGACACTCAGCCTCTGTACGCCATGAGACTCTCAGGAAGAAGATGTC 1800
QY 1801 CTGAACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGGACATCCTGGCTGAAGCAGAC 1860
Db 1801 CTGAACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGGACATCCTGGCTGAAGCAGAC 1860
QY 1861 ATGAGTGGGTGCTCAAACTCTACTACTCTCTCCAGGACAGGACAGCTGTACTTTGTG 1920
Db 1861 ATGAGTGGGTGCTCAAACTCTACTACTCTCTCCAGGACAGGACAGCTGTACTTTGTG 1920
QY 1921 ATGGACTACATACACAGGGGGATATGATGAGCCTGCTGATCAGGATGAGGCTCTCCCT 1980
Db 1921 ATGGACTACATACACAGGGGGATATGATGAGCCTGCTGATCAGGATGAGGCTCTCCCT 1980
QY 1981 GAGCACTTGGCCCGCTTCTACATTCAGAGTTGACCCCTGGCCATTGAAAGTGTCCACAAG 2040
Db 1981 GAGCACTTGGCCCGCTTCTACATTCAGAGTTGACCCCTGGCCATTGAAAGTGTCCACAAG 2040
QY 2041 ATGGCCTTATCCACCGGGACATCAAGCTGACAGTTCATCTACATCTGATGATGAT 2100
Db 2041 ATGGCCTTATCCACCGGGACATCAAGCTGACAGTTCATCTACATCTGATGATGAT 2100
QY 2101 ATTAAGCTGACAGATTTTGGCCTCTGCACTGATTCAGTGGACTCACAATTCCAAGTAC 2160
Db 2101 ATTAAGCTGACAGATTTTGGCCTCTGCACTGATTCAGTGGACTCACAATTCCAAGTAC 2160
QY 2161 TACCAGAAAGGAACCATGACAGGACAGCAGTGGAGCCCTGGAGCAGGCGCAGAGCAG 2220
Db 2161 TACCAGAAAGGAACCATGACAGGACAGCAGTGGAGCCCTGGAGCAGGCGCAGAGCAG 2220
QY 2221 GTTCCAACTGCTGCTGGAGACAGGTTAAGACCCCTGGAGCAGGCGCAGAGCAG 2280
Db 2221 GTTCCAACTGCTGCTGGAGACAGGTTAAGACCCCTGGAGCAGGCGCAGAGCAG 2280
QY 2281 CACAGAGGTGCTGGCAGATTCCTTGTGCGGACACCAATATACATCGCTCCGAGGTG 2340
Db 2281 CACAGAGGTGCTGGCAGATTCCTTGTGCGGACACCAATATACATCGCTCCGAGGTG 2340
QY 2341 CTTCTCCGAAAGGATACAGCAGCTCTGTGACTGGTGGAGCGTGGTGTGATTCCTTT 2400
Db 2341 CTTCTCCGAAAGGATACAGCAGCTCTGTGACTGGTGGAGCGTGGTGTGATTCCTTT 2400
QY 2401 GAGATGCTGTTGGCAGCGCTTCTTGGCCCGCCACCCACAGAGCAGCAGTGAAG 2460
Db 2401 GAGATGCTGTTGGCAGCGCTTCTTGGCCCGCCACCCACAGAGCAGCAGTGAAG 2460
QY 2461 GTGATCAACTGGGAGCAGCAGTGCATATCCCTACGAGGTGAGGCTCAGCGCTCAGGCC 2520
Db 2461 GTGATCAACTGGGAGCAGCAGTGCATATCCCTACGAGGTGAGGCTCAGCGCTCAGGCC 2520
QY 2521 CGAGACCTCATCAGAAAGTGTGCTGCGCGGTGACTCCCGCTGGGAGGATGGGCA 2580
Db 2521 CGAGACCTCATCAGAAAGTGTGCTGCGCGGTGACTCCCGCTGGGAGGATGGGCA 2580
QY 2581 GATGACCTCAAGGCACACCCGTTCTTCAACACCATTCGACTTTTCCCGTGCATCCGAAG 2640
Db 2581 GATGACCTCAAGGCACACCCGTTCTTCAACACCATTCGACTTTTCCCGTGCATCCGAAG 2640
QY 2641 CAGGCTGCACCTACGCTCCCAACCATCAGCCACCCATGGACACCTTCCAAATTTTCACCG 2700
Db 2641 CAGGCTGCACCTACGCTCCCAACCATCAGCCACCCATGGACACCTTCCAAATTTTCACCG 2700
QY 2701 GTGATGAAGAAGCCCTTGGCAGCAGGCTCAGCGAGAGGAGCCCAAGCCCTGGGACAG 2760
Db 2701 GTGATGAAGAAGCCCTTGGCAGCAGGCTCAGCGAGAGGAGCCCAAGCCCTGGGACAG 2760
QY 2761 CTGGCCTCCCGCAGCAGCAAGCATCCAGAGCAGCCTTCTATGATTCACCTTCGCGAGG 2820
Db 2761 CTGGCCTCCCGCAGCAGCAAGCATCCAGAGCAGCCTTCTATGATTCACCTTCGCGAGG 2820

QY 2821 TTCTTCGATGACAACGGCTATCCCTTCGGTGCCTGAAAGCCCTCAGAGCCGCGAGAGT 2880
Db 2821 TTCTTCGATGACAACGGCTATCCCTTCGGTGCCTGAAAGCCCTCAGAGCCGCGAGAGT 2880
QY 2881 GCAGACCCAGGGATGCGGACTTGAAGGTGCGGCCGAGGGCTGCCAGCCGGTGTACGTG 2940
Db 2881 GCAGACCCAGGGATGCGGACTTGAAGGTGCGGCCGAGGGCTGCCAGCCGGTGTACGTG 2940
QY 2941 TAAGCCTCAGTTAACCACAACCTCGAGGAACCCAAAATGAGATTCTTTTCAGAGACAA 3000
Db 2941 TAAGCCTCAGTTAACCACAACCTCGAGGAACCCAAAATGAGATTCTTTTCAGAGACAA 3000
QY 3001 ACTCAAGCTTAGGAATCCTTCAATTTTCTGTAATGGGCAACAGGAGAGCTCAAC 3060
Db 3001 ACTCAAGCTTAGGAATCCTTCAATTTTCTGTAATGGGCAACAGGAGAGCTCAAC 3060
QY 3061 ATGATTTCAAAATAGCCCTCTGAGGACCTTCACTGCATTAACACAGTATTTTAAAAA 3120
Db 3061 ATGATTTCAAAATAGCCCTCTGAGGACCTTCACTGCATTAACACAGTATTTTAAAAA 3120
QY 3121 TTAGTACAGTATGGAAGAGCAGCTTATTTTGGGG 3155
Db 3121 TTAGTACAGTATGGAAGAGCAGCTTATTTTGGGG 3155

RESULT 2

AAZ51507

ID AAZ51507 standard; DNA; 3155 BP.

XX AC AAZ51507;

XX AC 21-JUN-2000 (first entry)

XX Mouse Lats2 (large tumour suppressor) DNA.

Mouse; Lats2; large tumour suppressor; cytostatic; vulnery;
cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
treatment; prevention; screening; cancer; skin; ovarian tumour;
soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
dysplasia; degenerative disorder; growth deficiency; physical trauma;
hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX Mus musculus.

XX Key Location/Qualifiers

FH 1..2943

FT CDS /*tag= a

FT /product= "Lats2 protein"

FT 970..1920

FT /tag= b

FT /note= "This region is erroneously repeated in the

FT mouse Lats2 DNA sequence shown in figure 14"

FT 1921..2960

FT /tag= c

FT /note= "This region is missing in the mouse Lats2

FT DNA sequence shown in figure 14"

XX WO200010602-A1.

XX 02-MAR-2000.

XX 18-AUG-1999; 99WO-US19068.

XX 18-AUG-1998; 98US-0096996.

XX 18-AUG-1998; 98US-0096997.

XX (UYUA) UNIV YALE.

XX Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

PI Trenchalk GS, Stewart RA;

XX

Db 1681 AAGTCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCCCTTTGGGAAAGTGTCCCTC 1740
 QY 1741 GCTTGTAAAGCTGGACACTCAGCTCTGTAGCCATCAAGACTCTCAGGAAGAGGATGTC 1800
 Db 1741 GCTTGTAAAGCTGGACACTCAGCTCTGTAGCCATCAAGACTCTCAGGAAGAGGATGTC 1800
 QY 1801 CTGAACCCGAATCAAGTGGCCCATGTCAGAGCTGAGAGGACATCCTGCTCAAGCAGAC 1860
 Db 1801 CTGAACCCGAATCAAGTGGCCCATGTCAGAGCTGAGAGGACATCCTGCTCAAGCAGAC 1860
 QY 1861 AATGAGTGGTGGTCAAACTCTACTACTCTCTTCCAGGACAAAGACAGCTGTACTTTGTG 1920
 Db 1861 AATGAGTGGTGGTCAAACTCTACTACTCTCTTCCAGGACAAAGACAGCTGTACTTTGTG 1920
 QY 1921 ATGGACTACATACACAGCGGGGATATGATGAGCTCTGATCAGGATGAGGTCTTCCCTC 1980
 Db 1921 ATGGACTACATACACAGCGGGGATATGATGAGCTCTGATCAGGATGAGGTCTTCCCTC 1980
 QY 1981 GAGCACTGGCCCGCTTCTACATTCAGAGTGTGACCTTGGCCATTTGAAAGTGTCCCAAG 2040
 Db 1981 GAGCACTGGCCCGCTTCTACATTCAGAGTGTGACCTTGGCCATTTGAAAGTGTCCCAAG 2040
 QY 2041 ATGGGCTTTTATCCACCGGGACATCAAGCTGACACATATCTATCAGCTGGATGTCTAT 2100
 Db 2041 ATGGGCTTTTATCCACCGGGACATCAAGCTGACACATATCTATCAGCTGGATGTCTAT 2100
 QY 2101 ATTAAGCTCACAGATTTTGGCCCTCTGCACTGGATTCAGGTGACACATTCACAGTAC 2160
 Db 2101 ATTAAGCTCACAGATTTTGGCCCTCTGCACTGGATTCAGGTGACACATTCACAGTAC 2160
 QY 2161 TACCAGAAAGGAACACATGACAGACAGCATGAGCCCGGTGACCTCTGGGACGAT 2220
 Db 2161 TACCAGAAAGGAACACATGACAGACAGCATGAGCCCGGTGACCTCTGGGACGAT 2220
 QY 2221 GTTTCGAATCTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGCGCAGAGCAG 2280
 Db 2221 GTTTCGAATCTGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGCGCAGAGCAG 2280
 QY 2281 CACCAGAGTGTGCTGACATCTCTGTCGGGACACAAATACATCCTCCGGAGGTG 2340
 Db 2281 CACCAGAGTGTGCTGACATCTCTGTCGGGACACAAATACATCCTCCGGAGGTG 2340
 QY 2341 CTTCTCCGCAAAAGGTGTACAGCAGCTCTGTGCTGTGGAGCGTGGTGTGATCTCTTT 2400
 Db 2341 CTTCTCCGCAAAAGGTGTACAGCAGCTCTGTGCTGTGGAGCGTGGTGTGATCTCTTT 2400
 QY 2401 GAGATGCTGTGGGACGCGCTTCTTGGCCCGCCACCCACAGACAGCAGCTGAAG 2460
 Db 2401 GAGATGCTGTGGGACGCGCTTCTTGGCCCGCCACCCACAGACAGCAGCTGAAG 2460
 QY 2461 GTGATCACTGGGAGACGCTGTGATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCC 2520
 Db 2461 GTGATCACTGGGAGACGCTGTGATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCC 2520
 QY 2521 CGAGACCTCATCAGCAAGCTGTGCTGCGCGGTGACTGCGCGCTGGCAGGATGGGCA 2580
 Db 2521 CGAGACCTCATCAGCAAGCTGTGCTGCGCGGTGACTGCGCGCTGGCAGGATGGGCA 2580
 QY 2581 GATGACCTCAAGGCACACCCGTTCTCAACACCATCGACTTTCCCGTGACATCCGAAG 2640
 Db 2581 GATGACCTCAAGGCACACCCGTTCTCAACACCATCGACTTTCCCGTGACATCCGAAG 2640
 QY 2641 CAGGTGTACCTAGTGTCCCAACCATCAGCCACCCATGACACCTCCCAATTTGACCCG 2700
 Db 2641 CAGGTGTACCTAGTGTCCCAACCATCAGCCACCCATGACACCTCCCAATTTGACCCG 2700
 QY 2701 GTGGATGAAGAAAGCCCTTGGCAGCAGGCGCAGGAGAGGCGCCAAAGCCTGGGACAG 2760
 Db 2701 GTGGATGAAGAAAGCCCTTGGCAGCAGGCGCAGGAGAGGCGCCAAAGCCTGGGACAG 2760
 QY 2761 CTGGGCTTCCCGCAGCAGCAAGCATCCAGACAGCGCTTCTATGAGTTCACCTTCCGACAG 2820
 Db 2761 CTGGGCTTCCCGCAGCAGCAAGCATCCAGACAGCGCTTCTATGAGTTCACCTTCCGACAG 2820

QY 2821 TTCTTCGATGACAAAGGCTATCCCTTCGGTGCCTGAGCCCTCAGAGCCCGCAGAGCT 2880
 Db 2821 TTCTTCGATGACAAAGGCTATCCCTTCGGTGCCTGAGCCCTCAGAGCCCGCAGAGCT 2880
 QY 2881 GCAGACCCAGGGGATGCGGACTTGGAGGTGCGCGCGAGGCTGCCAGCGGTGTACGTG 2940
 Db 2881 GCAGACCCAGGGGATGCGGACTTGGAGGTGCGCGCGAGGCTGCCAGCGGTGTACGTG 2940
 QY 2941 TAAGCCTCAGTTAACCAACAACCTCGAGGAACCCAAATGAGATTTCTTTTCAGAAGACAA 3000
 Db 2941 TAAGCCTCAGTTAACCAACAACCTCGAGGAACCCAAATGAGATTTCTTTTCAGAAGACAA 3000
 QY 3001 ACTCAAGCTTAGGAATCCTTCAATTTTCTGTTAAATGGCAACAGAGAGTCAAC 3060
 Db 3001 ACTCAAGCTTAGGAATCCTTCAATTTTCTGTTAAATGGCAACAGAGAGTCAAC 3060
 QY 3061 ATGATTTCAAAATAGCCCTCTGAGGACCTTCACTGCAATTAACACAGTATTTTAAAAAA 3120
 Db 3061 ATGATTTCAAAATAGCCCTCTGAGGACCTTCACTGCAATTAACACAGTATTTTAAAAAA 3120
 QY 3121 TTAGTACAGTATGGAAGAGCAGCTTATTTTGGGG 3155
 Db 3121 TTAGTACAGTATGGAAGAGCAGCTTATTTTGGGG 3155

RESULT 3

AA87397

ID AA87397 standard; cDNA; 5276 BP.

XX AA87397;

XX AC

XX DT 08-OCT-1999 (first entry)

XX DE Human WART2 cDNA.

XX XX

KW WART2; hWART2; WART orthologue; human; signal transduction;

XX protein kinase; cancer; tumour; diagnosis; therapy; ss.

XX OS Homo sapiens.

XX XX

Key Location/Qualifiers

FT 5'UTR 1..374

FT CDS /*tag= a

FT 3'UTR 375..3641

FT /*tag= b

FT /*tag= c

PN W09937787-A2.

XX XX

PD 29-JUL-1999.

PF 20-JAN-1999; 99WO-US01145.

PR 21-JAN-1998; 98US-0072023.

PA (SUGEN-) SUGEN INC.

XX XX

PI Flanagan P, Plowman GD;

XX XX

DR WPI; 1999-458698/38.

XX XX

PT New nucleic acid encoding human orthologs of Drosophila WART

XX proteins, used to identify specific modulators for treating cancer

XX or for diagnosis

PS Claim 1; Page 122-123; 137pp; English.

XX XX

CC This is the nucleotide sequence of a cDNA clone coding for a human
 CC orthologue, i.e. hWART2 (see AAY06527), of Drosophila non-receptor
 CC serine/threonine kinase WART. hWART2 cDNA was isolated from a

human bone marrow cDNA library using a PCR fragment of hWART1 as probe. hWART2 is consistently expressed in human tumour cells lines, except for most of the colon cancer lines examined. This overexpression in tumour cells versus normal tissues may provide a target for oncology drug development. Nucleic acids encoding full-length hWART2 and hWART2 polypeptides lacking one or more of amino acid segments 1-33, 43-139, 342-466, 467-480, 514-518, and 974-1048, or lacking one or more of the N-terminal domain, catalytic domain, or C-terminal domain are claimed, as well as hWART1 sequences (AAH87396), hWART2 and hWART1 polypeptides, antibodies, a method for identifying modulators of hWART function, and use of such modulator compounds to treat an abnormal condition involving hWART signal transduction, especially cancer. Probes for detection of hWART nucleic acids are also claimed.

Sequence 5276 BP; 1348 A; 1407 C; 1345 G; 1176 T; 0 other;

Query Match 55.5%; Score 1749.6; DB 20; Length 5276;
Best Local Similarity 77.4%; Pred. No. 0;
Matches 2389; Conservative 0; Mismatches 544; Indels 153; Gaps 16;

QY 1 ATGAGACCCACCCGAGTTGGACCTTATCAAAAAGCTCTCAGGAAATCCGATATCC 60
DB 564 ATGAGACCCACCCGAGTTGGACCTTATCAAAAAGCTCTCAGGAAATCCGATATCC 623
QY 61 CTCCTGCTTTTCCCAACAGTCAAGCAGTTCGGCAGCTGCGAGGTGAACCGCAGATG 120
DB 624 TTGTGCTTTTCTAATGAATCGGCACCTCTGAGCTGCAAGAGTGAACCGCAATG 683
QY 121 CTTGAGAGTTGGTGAATCGCGGCTGTGACAGGAGATGGCTGGCAGAGCGCTCAGCAG 180
DB 684 CTCAGGAATGTTGTAACGAGTGTGACAGGAGTGGCTGGCAGAGCGCTCAGCAG 180
QY 181 ACGGCAGTAGGATGATCAAGCTGCTTGGAGTACATCAGTAAGTGGCTACCTGGAC 240
DB 744 ACTGGCAGGAGGATGATGAGCGCCGCTGGAGTACATCAGCAAGTGGCTACCTGGAC 803
QY 241 CCCAGGAATGACAGATTTGCGAGTTCATCAAGAGAGCTCCCAAGAAAGGCTGGCG 300
DB 804 CCGAGGAATGACAGATTTGCGAGTTCATCAAGAGAGCTCCCAAGAAAGGCTGGCG 863
QY 301 TCCACCCCGGTGACGCGGCGGAGTTTCGAGGACACAGGAGAGCACTCCCATCTAC 360
DB 864 CCAACCCCGGTGACGCGGCGGAGTTTCGAGGAGAGCACTCCCATCTAC 923
QY 361 CACAGCTGGGTGTGCAAACTACGAG-----GGCCCGCGCGCA 399
DB 924 CACAGCTGGGTGTGCAAACTACGAG-----GGCCCGCGCGCGCA 399
QY 400 CTGAGGAGATGCGCGGCAATATTTAGACTTTCTTCCCTGGAGCGCGGCGCGCAC 459
DB 984 CTGAGGAGATGCGCGGCGGCTACGTGGACTACCTTTTCCCGGAGTGGCGCGCGG 1043
QY 460 CAGGTGCGGAGCTCAGCAGATCTCCCAAGAGGTACAGCAGAGAGTAGGCAAGT 519
DB 1044 CCGCGCAACAG-----CACCAGACCCCAAGGCTTACGCTGCCAGCTGAGGCAAGCA 1100
QY 520 GCG-----CACTTTCGCGGCACACACTATGCTGCTGCTCATCTACT-----ATCG 564
DB 1101 GGGGCACACTTCCCGCTGCGAGGCGCGCATACGCGCGCGCGCGCGCGCGCTGG 1160
QY 565 GAGCAGTCTGGGTATGGGTGCGAGCGAGTTCTCTTCCAGAACAGAGCGCGCGCGAG 624
DB 1161 GAACCCCTGGGTACGAGGTGCGAGCGCGCGCATACGCGCGCGCGCGCGCGCGCTG 1220
QY 625 GCC-----TATTCCAGACTGGCC-----AAGGCCAGGTTGGCCCTTC-----CGCAGCTTC 672
DB 1221 ACCGGGGTTTACGCGAGCTGCGCGAGAGGCGCGAGGAGGAGCGCGCGCGCGCTC 1280
QY 673 ACCTTCTTCCGCGAGTGGGTGTACACTGCTTCGCGACCAAGCGCGCGCTACCCCA 732
DB 1281 GCTTTCACCGCTTCCGCGCGCGGCTTACGTGCGCGCACCCACACAGAGCGCGCT 1340

QY 733 CCTGGGCGCCACCCATTACATCTGTGGCAGCCCGGGTCCACAGCTTTACTGGGGAAGC 792
DB 1341 CCGCGGCGCCACCCAGCTGCTGGCTCCCGCAGCAGAGTGTTCGCGAGGAGAC 1400
QY 793 TCTGCACAGGCTGTGCTGGCAGCCCTCCAGAACAGCTCAATGTCTGTACGAGCTG 852
DB 1401 CCGCGCAGAGCTGCTCACTCCCTCCGCGAAGACAGCTCAAGCTGGACCTGTATGAATTG 1460
QY 853 GG-----CTCCACAGCTCCCTGGTCTCAGCTCCATGGCAGCCGCGAGCTGCTCGAG 906
DB 1461 GGCAGCAGCTCGCTCCAGAGTGGCGGCTGCCACCTGGCGCGCGGAGCTCCCTCGAG 1520
QY 907 AAGCAGGCTGTAGAGCCTCGCGG-----CGCATGTGGCTTTTCGGGCTGGC-----954
DB 1521 AAGCGCGGCTGGAGCGCGCGCGCGCGCGCTTCGGCTTCGGCTGAGTGGCGCAGTG 1580
QY 955 CCCAGCAGGACCAACTCTTCAACAACCCACA-----CT 990
DB 1581 CCCAGCAGGACCAACTCTTCAACAGCAGCAGCGCGCGCGCTTCGGCTTCGCAAGGCC 1640
QY 991 GAGCCCTCACTGCCCGCGCGCGCGCGCGCGCGCTGAGCGCGCTGAGCGCGCGCACATCTC 1050
DB 1641 GAGCCCTCCCTGCGCGCGCGCGCGCGCGCTGAGCGCGCTGAGCGCGCGCACATCTC 1700
QY 1051 CTTGTGAAGAGCTGTGCTGTGCGCGCGCGCGCGCGCGCTGAGCGCGCTGAGCGCGCG 1110
DB 1701 CCGGTGAAGAGCTGTGCTGTGCTGAGCGCGCGCGCGCGCGCTGAGCGCGCTGAGCGCG 1760
QY 1111 CCCGCTTGGTGGCTGGC-----CCACAGCAGCTTGGCAGCTGAG 1149
DB 1761 CCGGCTTGGTGGCTGGC-----GGCGCGCGCGCGCGCGCGCTGAGCGCG 1820
QY 1150 AGCTGAGAGCAAGGA-----GGCGCGCGCGCGCGCGCGCTGAGCGCG 1191
DB 1821 GGTTCGACCGCAAGGAGGAGCAGTGGCTGGCGCTGGCGCGCGCGCTTCCGCGTG 1880
QY 1192 GATGTGGACTATCGCGCTCCGAGCGCGCTGCGCGCGCGCTGAGCGCGCTGAGCGCGCTG 1251
DB 1881 GAGCTGAGTACGAGCGCGCGCGCGCGCGCTGAGCGCGCTGAGCGCGCTGAGCGCGCTG 1940
QY 1252 CTGCTGCCAGTAAGTCTGAGCAGTACAGCTGGAGCTGGAGCGCTGTGACCGAGCTGTG 1311
DB 1941 CTGCTGCCAGTAAGTCTGAGCAGTAC-----GACCTGGACAGCTGTGCGCGCATG 1994
QY 1312 CAGCAGAGTCTGCGAGGGG-----CACTGATCTAGAGCGGAGTACAGCGCGCGCGCG 1368
DB 1995 GAGCAGAGCTCGTGGCGCGCGCGCGCGCGCGCTGAGCGCGCTGAGCGCGCTGAGCGCG 2054
QY 1369 GCGAAGGAGCAAAAGCTGGCAGAGCAAAAGAGATTCAGACTCCCGCGCTGCGCTGTC 1428
DB 2055 GCGAAGGAGCAAAAGCTGGCAGAGCAAAAGAGATTCAGACTCCCGCGCTGCGCTGTC 2114
QY 1429 GCGAAGTACAGAGATGAGAGAGAGAGAGTCTGCGATCAAGAGTACTCCCTTAT 1488
DB 2115 GCGAAGTACAGAGATGAGAGAGAGAGAGTCTGCGATCAAGAGTACTCCCTTAT 2174
QY 1489 GCTTCAAAATCTTCATGAGCAACACCTGAGAGATGCTATCAAAAGCTTACCAGCAAG 1548
DB 2175 GCTTAAAGTCTTCATGAGCAGCAGTGGAGATGCTATCAAAAGCTTACCAGCAAG 2234
QY 1549 GTCAGCGGAGCTTACAGCTGGAGAGAGAGTGGCGAGAGTGGCGCTGTGAGCGCGAG 1608
DB 2235 GTTAAACCGAGGCTGCGAGCTGGAGAGAGAGTGGCGAGAGTGGCGCTGTGAGCGCTGAG 2294
QY 1609 CAGGAGAGATGAGAGAGTCTCTACAGAGAGAGTCTAACTACACCGCGCTGAGAGAG 1668
DB 2295 CAGGAGAGATGAGAGAGTCTCTACAGAGAGAGTCTAACTACACCGCGCTGAGAGAG 2354
QY 1669 GCGAAGTACAGAGTCCATGTTGTGAAATCAAGAGTCTAGGCGATCGGTGCTTTGG 1728
DB 2355 GCGAAGTACAGAGTCTATGTTGTCAAGTCAAAAGCTTGGGATCGGTGCTTTGGA 2414
QY 1729 GAAGTGTGCTGCTTGAAGCTGGACACTCAGCTCTGTACCGCATGAGAGTCTCAGG 1788

Db 2415 GAAGTGTGCTTCTTGTAGGTGGACATCAAGCCCTTACGCCATGAAGACCCCTAAAGG 2474
 QY 1789 AAGAAGGATGCTCTGAACCGGATCAAGTGGCCCATGTCAAGCTCAGAGGACATCCTG 1848
 Db 2475 AAAAGAGTGTCTTGAACCGGATCAAGTGGCCCATGTCAAGCTCAGAGGACATCCTG 2534
 QY 1849 GCTGAAGCAGACAATGAGTGGGTGCTCAAACTCTACTACTCTCTCTCAAGCAGGACAGC 1908
 Db 2535 GCCGAGGACAGACAATGAGTGGGTGCTCAAACTCTACTACTCTCTCAAGCAGGACAGC 2594
 QY 1909 CTGTACTTGTGATGACTACATACACAGCGGGGATATGATGAGCCTGTCTGATCAGGATG 1968
 Db 2595 CTGTACTTGTGATGACTACATACACAGCGGGGATATGATGAGCCTGTCTGATCAGGATG 2654
 QY 1969 GAGGTCTTCCCTGAGCAGCCTGCGCTTCTACATTCAGAGTGTACCCCTGCGCATTTGAA 2028
 Db 2655 GAGGTCTTCCCTGAGCAGCCTGCGCTTCTACATTCAGAGTGTACCCCTGCGCATTTGAG 2714
 QY 2029 AGTGTCCACAAGATGGCTTTATCCACGGGACATCAAGCCTGACAACTACTCTATCTGAC 2088
 Db 2715 AGTGTCCACAAGATGGCTTTATCCACGGGACATCAAGCCTGACAACTACTCTATCTGAC 2774
 QY 2089 CTGGATGGTCAATTAAGCTGACAGATTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2148
 Db 2775 CTGGATGGTCAATTAAGCTGACAGATTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTG 2834
 QY 2149 AATTCCAAGTACTACACAGAGGACACATGACAGCAGCAGCAGCAGCAGCAGCAGCAGC 2208
 Db 2835 AATTCCAAGTACTACACAGAGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2894
 QY 2209 CTCTGGGAGGATGCTTTTCAACTGCTGCTGGAGCAGGTTAAAGACCCCTGGAGCAGAG 2268
 Db 2895 CTCTGGGAGGATGCTTTTCAACTGCTGCTGGGAGCAGGTTAAAGACCCCTGGAGCAGAG 2954
 QY 2269 GCGCAGACAGCAGCAGAGGCTGCTGGCAGATTTCTTTGCTGGGACACCAATATACATC 2328
 Db 2955 GCGCAGACAGCAGCAGAGGCTGCTGGCAGATTTCTTTGCTGGGAGCAGCAGCAGCAGC 3014
 QY 2329 GCTCCGGAGGCTTCTCCGAAAGGTTACACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 2388
 Db 3015 GCACCGAGGCTCTCTCCGAAAGGTTACACGAGCTGCTGCTGCTGCTGCTGCTGCTGCTG 3074
 QY 2389 GTGATTTCTTTTGTAGATGCTGGTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2448
 Db 3075 GTGATTTCTTTTGTAGATGCTGGTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3134
 QY 2449 AGCAGCTGAAGTGTCAACTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2508
 Db 3135 ACCCAGCTGAAGTGTCAACTGGGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3194
 QY 2509 AGCGCTGAGGCGGAGACCTCATCAGAGCTGTGCTGCGCGGCTGACTGCGCGCTGGGC 2568
 Db 3195 AGCGCTGAGGCGGAGACCTCATCAGAGCTGTGCTGCGCGGCTGACTGCGCGCTGGGC 3254
 QY 2569 AGGATGGGGAGATGACTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2628
 Db 3255 CGGAATGGGGAGATGACTCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 3314
 QY 2629 GACATCCGAAAGCAGGCTGACCTTACCTGCTCCACCATCAGCAGCAGCAGCAGCAGCAGC 2688
 Db 3315 GACATCCGAAAGCAGGCTGACCTTACCTGCTCCACCATCAGCAGCAGCAGCAGCAGCAGC 3374
 QY 2689 AATTTTGAACCGGTGGATGAAGAACCCCTTGGCAGGAGGAGCAGCAGCAGCAGCAGCAGC 2748
 Db 3375 AATTTTGAACCGGTGGATGAAGAACCCCTTGGCAGGAGGAGCAGCAGCAGCAGCAGCAGC 3434
 QY 2749 GCTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2808
 Db 3435 GCTGGGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 3494
 QY 2809 ACCTTCGCGAGGTTCTTCGATGACAGGCTATCCCTTCCGGTGGCCGAGCCCTCAGC 2866

Db 3495 ACCTTCCGAAGGTCTTTTGTATGACAATGCTACCCCTTTTCGATGCGCAAGACCTTCAGGA 3554
 QY 2867 -----ACCCGCGAGAGTGCAGACCCAGGGGATCGGAGCTTGAAGGTGCGGCC 2916
 Db 3555 GCAGAGCTTCACAGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGGTGGATCAGACT 3614
 QY 2917 GAGGCTGCCAGCCGCTGTACGTGA 2942
 Db 3615 GAAGCTGCCAGCTCTGTACGTGA 3640
 RESULT 4
 AAA59129
 ID AAA59129 standard; DNA; 5486 BP.
 XX
 AC AAA59129;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE DNA encoding a tumour suppressor protein hGHITS1.
 KW Human: growth hormone inhibited tumour suppressor protein; hGHITS1;
 KW antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 KW diabetic nephropathy; cardiopathy; tumour; breast cancer;
 KW renal adenocarcinoma; colorectal cancer; leukaemia; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 387..3653
 FT /tag= a
 FT /product= "tumour suppressor protein"
 XX
 PN EP1022333-AL.
 XX
 PD 26-JUL-2000.
 XX
 XX
 PF 07-OCT-1999; 99EP-0119199.
 XX
 PR 25-JAN-1999; 99JP-0016223.
 XX
 PA (JCRP-) JCR PHARM CO LTD.
 XX
 PI Koga J, Kono K, Zolotaryov FN;
 XX
 DR WPI: 2000-516013/47.
 DR P-PSDB; AAB07633.
 XX
 PT New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 PT diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 PT nephropathy or cardiopathy
 XX
 PS Claim 1; Page 18-28; 59pp; English.
 XX
 CC The present sequence encodes a human growth hormone inhibited tumour
 CC suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 CC antineoplastic pharmaceutical preparation. Probes for the hGHITS
 CC DNA sequences can be used in diagnostic pharmaceutical preparations.
 CC The diagnostic pharmaceutical preparations can be used for examining
 CC expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 CC angiopathy, diabetic nephropathy or cardiopathy, or in malignant
 CC tumours including breast cancer, renal adenocarcinoma, colorectal
 CC cancer, and leukaemia. Antibodies against the proteins can be used in
 CC a diagnostic pharmaceutical preparation for examining expression of
 CC a tumour suppressor gene.
 XX
 SQ Sequence 5486 BP; 1378 A; 1452 C; 1390 G; 1266 T; 0 other;
 Query Match 55.5%; Score 1749.6; DB 21; Length 5486;
 Best Local Similarity 77.4%; Pred. No. 0;
 Matches 2389; Conservative 0; Mismatches 544; Indels 153; Gaps 16;
 QY 1 ATGAGAGCCACCCCGAAGTTTGGACCTTATCAAAAAGCTCTCAGGGAATCCGATATTC 60

||||| 576 ATGAGAGCCACCCAAAGTTGCGACCTTATCAGAAAGCCTTGAGGAAATCAGATATTC 635
QY 61 CTCCTGCTTTTCCCAACGAGTCAGGACTTCGGAGCTGCGAGAGTGAACCGCAGATG 120
Db 636 TTGTTGCTTTTGTCTAATCAATCGGCGACTCTGCGAGTGCAGAGTGAACCGCAATG 695
QY 121 CTTGAGGAGTTGGTGAATCGGCGATGTGACAGAGAGATGGCTGGCAGAGCGCTACGAG 180
Db 696 CTCAGGAACTGGTGAACGAGATGCGACAGGAGATGGCTGGCGGAGCTCTCAAGCAG 755
QY 181 ACGGCGAGTAGAGATCAAGAGTGGCTGCTTGGAGTACATCAGTAAGATGGCTACCTGGAC 240
Db 756 ACTGGCAGCAGGAGATCGAGCGCCCTTGGAGTACATCAGCAAGATGGCTACCTGGAC 815
QY 241 CCGAGGATGAGCAGATGTGCGAGTCTCAAGCAGACTCCCGCAGGAAAGGCTGCGG 300
Db 816 CCGAGGATGAGCAGATGTGCGGCTCATTAAGCAGACTCCCGCAGGAAAGGCTGATG 875
QY 301 TCCACCCGGGTGACTCGGCGGCCAGTTTCGAGGACACAGGGAAGCAGTCCCATCTAC 360
Db 876 CCAACCCAGTGACCGGAGGCCAGCTTCGAAGGAACCGCGATTCTGCGCTCTAC 935
QY 361 CACGAGTGGGTGCAAACTACAG-----GGCCCCCGCGCA 399
Db 936 CACGAGTGGGGTACCCCTACAGAGGGCCCAAGCTTCGCGCTGAGCGGCCCCACGGCG 995
QY 400 CTGAGGAGATGCCCGCGCAATATTAGACTTCTTCCCTGGAGCGGAGCGGCGAC 459
Db 996 CTGAGGAGATGCCCGCGCTAGCTGACTACCTTTCCCGGAGTGGCGGCCCGCGG 1055
QY 460 CACGCTGCCAGGCTCACAGATCTCCCAAGGATACACAGCAGCAGTACAGCCAAAGT 519
Db 1056 CCGGCCACCAAG-----CACACAGCACCACCAAGGGCTACGGTGCAGCGTAGAGCAGCA 1112
QY 520 GCG-----CACTTTCGGGACACACTATGTGCTGTGCTACT-----ATCG 564
Db 1113 GGGCACACTTCCCGCTCAGGGCGGCTACGGGCGGCGCCACCTGCTGGTGGCTGGG 1172
QY 555 GAGCAGTCTGGGTATGGGTGACGCGAGTTCTCTTCCAGAAAGAGCGCCACCAAGAT 624
Db 1173 GAACCCCTGGGCTACGGAGTACGCGAGCGCCCTCTTCCAGAGCAGCGCCCGCGAG 1232
QY 625 GCG-----TATTTCCAGCATGCGC---AAGCCCGAGGTGGCGCTCC---CGCCAGCCTC 672
Db 1233 ACCGGGGTTACGGCAGCTGCGCAGAGGGCCAGGGAGCGCGCAGCGCGCGCTC 1292
QY 673 ACCTTTCCTGCGCATGCTGGGCTGTACACTGCTCGCACACAGCGCGGCTACCCCA 732
Db 1293 GCTTTCACCCCTGCGCGCGGCTCTAGTGGCGCACCCACACAGCAGCGCGGT 1352
QY 733 CTTGGGCGCCACCATTTACATGTGTTGGCACCGGGGTTCCACGTTTACTGGCGAAAGC 792
Db 1353 CCGCGCGCCACCAAGCTGCTGCTGGCTCCCGAGCGAGGTGTCGCGAGCAGC 1412
QY 793 TCTCACAGCTGTGCTGCGCAGCTCCAGCAAGCCTCAATGCTGACTGTAGCAGCTG 852
Db 1413 CCCCCGAGAGCCTGCTACTCTCCCGGAACAGCGCTCAACGTGGAGCCTGTATGAATTG 1472
QY 853 GG-----CTCCAGGTGGCTGTGCTGACCTCCACTGGCAGCGCGGCTCGCTGCAG 906
Db 1473 GCGAGCCTTCCGTCAGAGTGGCGGCTGCCACCTTGSCCGCGGAGTCCCTGCAG 1532
QY 907 AAGCAGGCTTAGAGCTCGCGC---CGCATGTGGCTTTTCGGGCTGGC-----954
Db 1533 AAGCGGGCTTGAGGCGCGCGCGCGCTGCGGCTTCCGGCTGACTGCCAGTG 1592
QY 955 CCGAGGAGCAACTCTCTCAACACCCCAAC-----CT 990
Db 1593 CCGAGGAGCAACTCTCTCAACAGCCACCGCGCGCGCGCTGCGGCTGGCAAGGCC 1652
QY 991 GAGCCTCTAGTCCGCGCCCAACAGGTACCGCGCTGACGCGCGCACACATCTCTAC 1050
|||||

Db 1653 GAGCCTCCCTGCGCGCCCAACACCGTGTGAGGGTGTACGGCGCGCACATCTTGAC 1712
QY 1051 CTTGTGAAGAGCGTGTGCTGTGGGCGCCGAGCCCGCAGACAGCGTGGGCGCCTCGCAC 1110
Db 1713 CCGGTGAAGAGCGTGTGCTGTGAGCGCGGAGCGCGTGTGGGCGCCTCGCAC 1772
QY 1111 CCGCCTGGTGGTGTGCGC-----CCACAGCAGCTGCGCACAG 1149
Db 1773 CCGCCTGGTGGTGTGCGCCTGCGCGCCCGCGCGCGCGCGCGCGCGCGCGCG 1832
QY 1150 AGCCTGGAGACCAAGG-----GGCAGCGCAGGCGCCACACCGCGTG 1191
Db 1833 GCTTGGAGCGCAAGGAGGAGCATGCCGTGGCGTGGCGCGCAGCGGCTTCCCGCTG 1892
QY 1192 GATGTGACTATGGCGGCTCGGAGCGAGTGGCCACCGGCTTCGTATPCCAAGACCTTG 1251
Db 1893 GACGTGAGTACGAGGCGCCAGACCGGAGTGGCGGCTCGCGCTACCGCAAGCAGCTG 1952
QY 1252 CTGCTGCCAGTAAAGTCTGAGCAGTACAGCGTGGACCTGGACAGCGCTGTGCACAGTGTG 1311
Db 1953 CTGCTGCGCAAGTCTGGAGCAGTAC-----GACCTGGACAGCGCTGTGCGCAGCATG 2006
QY 1312 CAGCAGAGTCTGCGAGGGG---CACTGATCTAGACGGGAGTGCACAGAGCCACCAAGGT 1368
Db 2007 GAGCAGAGCTCTGCGGGCCCAACGAGCGCGGAGCGCGGAGAGCGCGCAAGC 2066
QY 1369 GCGAAGGAGACAAAGCTGCGAGAGCAAAAGAGATTCAGACCTCCCGGTCGCTGTC 1428
Db 2067 GCGAAGGAGGAGCAAGCGCGGAAAGGATAAAGAGATTCAGACCTCTCCCGTTCGCGTC 2126
QY 1429 CGCAAGATAGCAGATGAAGAGAGAGAGTCTCGCATCAAGAGTCTACCTCCCTTAT 1488
Db 2127 CGCAAAACAGCAGAGACGAGAGAGAGAGTACGCGCATCAAGAGTCTACTCGCCATAC 2186
QY 1489 GCCTTCAATCTTCATGGAGCAACAGTGTGAGAAATGTATCAAAACCTACCAAGAGAG 1548
Db 2187 GCCTTAAAGTCTTCTATGGAGCAGCAGTGGAGAAATGTATCAAAACCTACCAAGAGAG 2246
QY 1549 GTCAGCGGAGGCTACAGCTGGAGCAAGTGGCAAGCTGGGCTCTGTGAGCGCGAG 1608
Db 2247 GTTAAACGGAGGCTGACGCTGGAGCAAGAAATGGCAAGCTGGAGTCTGTGAGAGTGA 2306
QY 1609 CAGGAGCAGATGAGGAGATCTCTTACAGAGAGTCTAACTACAAACCGCTGAAGAGG 1668
Db 2307 CAGGAGCAGATGCGGAAGATCTCTTACCAGAAAGTCTAAATACACAGGTAAAGAGG 2366
QY 1669 GCCAAGATGCAAGTCCATGTTGTGAATAACAGACTCTAGGCAATCGGTCGCTTTGGG 1728
Db 2367 GCCAAGATGCAAGTCTATGTTGTCAAGATCAAAACCTGGGGATCGGTCGCTTTGGA 2426
QY 1729 GAAGTGTGCTCGCTTGTAAAGTGGAGACTCAGCGTCTGTAGCCATGAAGACTCTCAGG 1788
Db 2427 GAAGTGTGCTTGTAAAGTGGAGACTCAGCGCTGTAGCCATGAAGACTCTTAAGG 2486
QY 1789 AGAAGATGTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTCTG 1848
Db 2487 AAAAGGATGTCTGAACCGGAATCAGTGGCCACGTCAAGCGCGAGAGGACATCTCTG 2546
QY 1849 GCTGAAGCAGACAAATGAGTGGGTGCTCAACTCTACTCTCTCCAGGAGAGAGCAGC 1908
Db 2547 GCGAGGAGACAAATGAGTGGGTGCTCAACTCTACTCTCTCCAGGAGAGAGCAGC 2606
QY 1909 CTGTACTTGTGATGAGTACATACAGCGGGGATATGATGAGCTGTCTGATCAGGATG 1968
Db 2607 CTGTACTTGTGATGAGTACATCTCTGTGGGAGATGATGAGCTGTCTGATCGGATG 2666
QY 1969 GAGTCTTCCCTGAGCACTGCGCGCTTCTACATTTGAGAGTTGACCTGCGCATTAAG 2028
Db 2667 GAGGTCTTCCCTGAGCAGCTGCGCGGTTCTACATCGCAGAGTCTACTTTGGCCATTAAG 2726
QY 2029 AGTGTCCCAAGATGGCTTTATCCACGGGAGATCAAGCCTGACCAACATCTCTATCAGC 2088
Db 2727 AGTGTCCCAAGATGGCTTCTATCCAGCAGACATCAAGCCTGATACATTTTATAGAT 2786

QY 2089 CTGGATGCTATTAAGCTGACAGATTTGGCTCTGCACCTGGATTGAGTGGGACTCAC 2148
 Db 2787 CTGGATGCTACATTAAGCTGACAGATTTGGCTCTGCACCTGGATTGAGTGGGACTCAC 2846
 QY 2149 AATTCCAAGTACTACAGAAAGGGAACACATGACAGAGGACAGATGGAGCCCGGTGAC 2208
 Db 2847 AATTCCAAGTACTACAGAAAGGGAACACATGACAGAGGACAGATGGAGCCCGGTGAC 2906
 QY 2209 CTCTGGGAGCTGTTTCCAACTGCTGCTGGAGACAGTTAAAGACCTGAGCAGAGG 2268
 Db 2907 CTCTGGGAGCTGTTTCCAACTGCTGCTGGAGACAGTTAAAGACCTGAGCAGAGG 2966
 QY 2269 GCGCAGAAAGCAGCAGAGAGTGCCTGGCACAATTTCTTGTGGGACACCAATATACATC 2328
 Db 2967 GCGCAGAAAGCAGCAGAGAGTGCCTGGCACAATTTCTTGTGGGACACCAATATACATC 3026
 QY 2329 GCTCGGAGGCTCTCTCGGAAAGGTTACAGCAGCTCTGTGACTGGTGGAGCGTCGGT 2388
 Db 3027 GCTCGGAGGCTCTCTCGGAAAGGTTACAGCAGCTCTGTGACTGGTGGAGCGTCGGT 3086
 QY 2389 GTGATCTCTTTGAGATGCTGCTGGGAGCGGCTTTCTTGGCCGCCACCCACAGAG 2448
 Db 3087 GTGATCTCTTTGAGATGCTGCTGGGAGCGGCTTTCTTGGCCGCCACCCACAGAG 3146
 QY 2449 AGCAGCTGAAGGATGATCAACTGGGAGAGCAGCTGTCATATCCCTACGAGGTGAGGCTC 2508
 Db 3147 AGCAGCTGAAGGATGATCAACTGGGAGAGCAGCTGTCATATCCCTACGAGGTGAGGCTC 3206
 QY 2509 AGCGTGAAGGCGGAGAGCTCTACAGAGCTGCTGCTGGGAGCTGCTGCTGGGAGCTG 2568
 Db 3207 AGCGTGAAGGCGGAGAGCTCTACAGAGCTGCTGCTGGGAGCTGCTGCTGGGAGCTG 3266
 QY 2569 AGGGATGGGAGAGTCAAGCAGCAGCAGCTTCTTCAACACCATGACTTTTCCGCT 2628
 Db 3267 AGGGATGGGAGAGTCAAGCAGCAGCAGCTTCTTCAACACCATGACTTTTCCGCT 3326
 QY 2629 GACATCCGAAGAGGCTGACCTTACGTCGCCACCATCAGCAGCAGCTGACACCTCC 2688
 Db 3327 GACATCCGAAGAGGCTGACCTTACGTCGCCACCATCAGCAGCAGCTGACACCTCC 3386
 QY 2689 AATTTGACCGGCTGATGAAGAAAGCCCTGGCAGAGGAGGAGGAGGAGGAGGAGG 2748
 Db 3387 AATTTGACCGGCTGATGAAGAAAGCCCTGGCAGAGGAGGAGGAGGAGGAGGAGGAGG 3446
 QY 2749 GCCTGGGACAGCTGCTGCTCCCGCAGCAGCAGCTTCTTCAACACCATGACTTTTCCGCT 2808
 Db 3447 GCCTGGGACAGCTGCTGCTCCCGCAGCAGCAGCTTCTTCAACACCATGACTTTTCCGCT 3506
 QY 2809 ACCTTCCGAGGCTTCTCGATGACAAAGGCTATCCCTTCCGCTGCCCGAGCCCTCAG-- 2866
 Db 3507 ACCTTCCGAGGCTTCTCGATGACAAAGGCTATCCCTTCCGCTGCCCGAGCCCTCAG-- 3566
 QY 2867 -----AGCCCGGAGAGGTCACAGCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 2916
 Db 3567 GCAGAGGCTTACAGGCTGAGAGCTCAGATTTAGAAAGCTCTGATCTGTGATGATGAGT 3626
 QY 2917 GAGGCTGCCAGCGGCTGACTGTA 2942
 Db 3627 GAGGCTGCCAGCGGCTGACTGTA 3652

RESULT 5
 AAA59130
 ID AAA59130 standard; DNA; 5486 BP.
 XX
 AC AAA59130;
 XX
 DT 07-NOV-2000 (first entry)
 XX
 DE DNA encoding a tumour suppressor protein hGHITS2.
 XX
 KW Human; growth hormone inhibited tumour suppressor protein; hGHITS;

antineoplastic; dwarfism; gigantism; acromegaly; angiodysplasia; angiodysplasia;
 diabetic nephropathy; cardiopathy; tumour; breast cancer;
 renal adenocarcinoma; colorectal cancer; leukaemia; ss.
 XX Homo sapiens.
 XX Key Location/Qualifiers
 FH CDS 387..3653
 FT /*tag= a
 FT /*product= "tumour suppressor protein"
 XX
 PN EP1022333-A1.
 XX 26-JUL-2000.
 PD 07-OCT-1999; 99EP-0119199.
 PF 25-JAN-1999; 99JP-0016223.
 XX (JCRP-) JCR PHARM CO LTD.
 PA Koga J, Kono K, Zolotaryov FN;
 PI WPI; 2000-516013/47.
 XX P-PSDB; AAB07664.
 DR New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 XX diagnosis of dwarfism, gigantism, acromegaly, angiodysplasia, diabetic
 XX nephropathy or cardiopathy
 XX Claim 1; Page 28-39; 59pp; English.
 XX The present sequence encodes a human growth hormone inhibited tumour
 XX suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 XX antineoplastic pharmaceutical preparation. Probes for the hGHITS
 XX DNA sequences can be used in diagnostic pharmaceutical preparations.
 XX The diagnosis of hGHITS genes in dwarfism, gigantism, acromegaly,
 XX angiodysplasia, diabetic nephropathy or cardiopathy, or in malignant
 XX tumours including breast cancer, renal adenocarcinoma, colorectal
 XX cancer, and leukaemia. Antibodies against the proteins can be used in
 XX a diagnostic pharmaceutical preparation for examining expression of
 XX a tumour suppressor gene.
 XX Sequence 5486 BP; 1380 A; 1451 C; 1388 G; 1267 T; 0 other;
 SQ
 Query Match 55.4%; Score 1748; DB 21; Length 5486;
 Best Local Similarity 77.4%; Pred. No. 0;
 Matches 2388; Conservative 0; Mismatches 545; Indels 153; Gaps 16;
 QY 1 ATGAGAGCAGCCCGCAAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCGATATTC 60
 Db 576 ATGAGAGCAGCCCGCAAGTTTGGACCTTATCAAAAGCTCTCAGGAAATCGATATTC 635
 QY 61 CTCCTGCCCTTTTGGCAACAGTACAGGCTTGGCAGCTGACAGAGTGAACCGGAGATG 120
 Db 636 TTGTTGCCCTTTTGGCAACAGTACAGGCTTGGCAGCTGACAGAGTGAACCGGAGATG 695
 QY 121 CTTAGGAGTGGTGAATGCGGCATGTACACAGGAGATGGTGGCAGAGCGCTCAGCAG 180
 Db 696 CTGCAGGAAGTGGTGAACGAGGATGCGACCAAGAGATGGTGGCAGAGCGCTCAGCAG 755
 QY 181 ACGGCGAGTGGAGTATCAAGCTGCTTGGAGTACATCAGTAACTACCTGGCTACCTGGAC 240
 Db 756 ACTGGCAGCAGAGATGAGGCGGCTTGGAGTACATCAGCAAGATGGCTTACCTGGAC 815
 QY 241 CCGAGGAATGAGCAGATTTGTGGAGTCTATCAAGCAGAGCTCCCGAGGAAGGCTCGGC 300
 Db 816 CCGAGGAATGAGCAGATTTGTGGAGTCTATCAAGCAGAGCTCCCGAGGAAGGCTCATG 875
 QY 301 TCCACCCCGGCTGCTGCGGCGGCTTTCAGGAGGAGCAGGAGGAGGAGGAGGAGGAGG 360
 Db 876 CCAACCCCGGCTGCTGCGGCGGCTTTCAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 935

QY 361 CACAGCTGGTGGTCAAACTACGAG-----GGCCCGCGCGCA 399
 DB 936 CACAGCTGAGGTACCCCTACAGGGCCCAAGCTTCGGCGGTGACGGCCCAACGGCG 995
 QY 400 CTGAGGAGATGCGCGCGCAATATTTAGACTTTCTTTCCCTGAGCGGAGCGGCAACC 459
 DB 996 CTGAGGAGATGCGCGCGCGCTAGCTGAGTACCTTTTCCCGGAGTCCGCGCCCGCGG 1055
 QY 460 CACGGTCCCGAGGTACACAGCATCTCCCAAGGGTACACAGCAGTAGAGCCAGT 519
 DB 1056 CCGCGCCACAG-----CACAGACCCACCCCAAGGGTACGGTGCAGCGTAGAGCGACGA 1112
 QY 520 GCG-----CACTTTCCGGGCACACACTATGCTGGTGTCTACT-----ATCG 564
 DB 1113 GGGGCACACTTCCCGCTGCGAGGGCGGCACTACGGGCGCGGCACCTGCTGGTGCCTGGG 1172
 QY 565 GAGCAGTCTGGGTATGGGTGACGGAGTTTCCTTCCTTCAGAAACAGAGCCACACAGAT 624
 DB 1173 GAACCCCTGGGCTACGGAGTGCAGCGAGCCCTCTCTTCAGAGACAGAGCGCGCGGAG 1232
 QY 625 GCG-----TATTCACAGCATGGCC---AAGGCCAGGTGGCCCTCC---CGCCAGCCTC 672
 DB 1233 ACCGGGGTTACGGCGAGCTGCCCCACAGAGGGCCAGGAGGAGCCGAGGCGCGGCGCTC 1292
 QY 673 ACCTTTCTGCCCCATGCTGGGCTGTACACTGCTCTGCGACACAGCGGGGGCTTACCCCA 732
 DB 1293 GCTTTCCACCCCTGCGCGCGGCTCTAGTGTGCGGCGGCAACACACAGAGCGCGGT 1352
 QY 733 CTTGGGCGCCACCATTTACATGTTGGGACACCGGGGTCCCGAGTTTACTTGGGGAAGC 792
 DB 1353 CCGTGGCGCCACCACTGCTGCTGGGCTCCCGGAGCAGAGTGTTCGCGAGGACGAC 1412
 QY 793 TCTCACAGGCTGTGCTGGCAGCTTCCAGGAACAGCTCAATGCTGACTGTACGAGCTG 852
 DB 1413 CCCCCGAGAGCTGTCTACTCTCCCGGAACAGCTCAACGTGGACCTGTATGAATG 1472
 QY 853 GG-----CTCCAGGTGCGCTGTCTGCACTCCACTGGCAGCCCGCGACTGCTGTCAG 906
 DB 1473 AGCAGCACTCCGTCAGAGTGTGCGGCTGCCACCTGCGCGCGCGGACTCCCTGCGAG 1532
 QY 907 AAGCAGGCTTAGAAGCTTCCGCGC---CGATGTGCTTTTGGGCTGGC-----954
 DB 1533 AAGCGGGCTTGGAGGCGCGCGCGCGCAGCTGCGCTTCCGGCTGACTGCCCGAGT 1592
 QY 955 CCCAGCAGGACCACTCTTCAACAAACCCACAAC-----CT 990
 DB 1593 CCCAGCAGGACCACTCTTCAACAGCCACAGCGCGCGCGCGCTGCGCTGGCAAGGCC 1652
 QY 991 GAGCCCTCACTGCGCGCGCCCAACACAGGTCAACCGCTGACGGCGGCACACACTCTTTCAC 1050
 DB 1653 GAGCCCTCCCTGCGCGCGCCCAACACAGGTGACGGCTGTACGGCGCGGCACACTTTCAC 1712
 QY 1051 CTTGTGAAGAGCTGTGCTGCTGCGCGCGAGCCCGCAGACAGCGCTGGGCGCTGCGAC 1110
 DB 1713 CCGTGAAGAGCTGTGCTGCTGAGGCGCGAGCGCGGAGCGGCTGTGGGCGCTGCGAC 1772
 QY 1111 CCGCGCTGGTGGCTGCGC-----CCACAGCACTGCGACCTGAG 1149
 DB 1773 CCGCGCTGGTGGCTGCGCGCGCTGCGCGCGCGCGCGCGCGCGCTGGCGGAG 1832
 QY 1150 AGCTTGAGACGAAGGA-----GGCGAGCGAGCGCCACACCGCGCTG 1191
 DB 1833 GGCTTGACGCCAAGGAGGAGCATGCTTGGCGTGGCGGCGAGCGCGCTTCCCGCTG 1892
 QY 1192 GATGTGAGCTATGCGGCTCCGAGCGAGGTGCGCCACCGCTTGGTATCCAAAGCACTTG 1251
 DB 1893 GACTGTGAGTACGAGGCGCGAGCGGAGGTGCGCGCTCCGCGCTACCGGGAAGCACTG 1952
 QY 1252 CTGCTGCCAGTATGCTGAGCAGTACAGCTGAGCTGAGACCTGTCACAGCTGTGCGAGTGTG 1311
 DB 1953 CTGCTGCCAGCAAGTGGAGCAGTAC-----GACCTGGACAGCTGTGCGGAGGCATG 2006

QY 1312 CACAGAGTCTCGAGGGGG---CACTCATCTAGACGGGAGTGAACAAGACCCACAAGGT 1368
 DB 2007 GAGCAGAGCTCTCCGTGCGGGCCCCAACAGCGCGGCGGCGAGCAAGCGCCCAAAAGC 2066
 QY 1369 GCGAAGGAGACAAAGCTGGCAGAGACAAAAGCAGATTCAGACCTCCCGGTCCTGTC 1428
 DB 2067 GCCAAGGGGACAAAGCGGAAAGGATAAAAGCAGATTCAGACCTCTCCCGTTCGGCTC 2126
 QY 1429 CGCAAGATAGCAGAGTGAAGAGAGAGTCTCGCATCAAGAGTTACTCCCTCCCTAT 1488
 DB 2127 CGCAAAACAGCAGAGAGAGAGAGAGTCAAGCATCAAGAGCTACTCCCATAC 2186
 QY 1489 GCCTTCAATTTTCATGGAGCAACAGCTGGAGAAATGTCAATCAAAACCTACACAGAG 1548
 DB 2187 GCCTTTAAGTTCTTCATGGAGCAGCAGTGGAGAAATGTCAATCAAAACCTACACAGAG 2246
 QY 1549 GTCAGCCGGAGCTACAGCTGGAGCAGAAATGCCAAAGCTGGGCTCTGTGAGGCGGAG 1608
 DB 2247 GTTAACCGGAGGCTGCGAGCTGGAGCAAGAAATGCCAAAGCTGGAGCTCTGTGAAGCTGAG 2306
 QY 1609 CAGGAGCAGATGAGGAAGATCTCTACCAAGAGAGTCTAACTACAACCGCTCGAAGG 1668
 DB 2307 CAGGAGCAGATGCGGAAGATCTCTACCAAGAGAGTCTAACTACAAGGTTAAAGAGG 2366
 QY 1669 GCCAAGATGGAACAGTCCATCTTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGG 1728
 DB 2367 GCCAAGATGGAACAGTCTATGTTGTCAAGATCAAAACCTGGGGATCGGTGCTTTGGA 2426
 QY 1729 GAAGTGTGCTGCTTGAAGCTGGAGCAGTCAACCTCTGTACGCCATCAAGACTCTCAGG 1788
 DB 2427 GAAGTGTGCTGCTTGAAGCTGGAGCAGTCAACCTCTGTACGCCATCAAGACTCTCAGG 2486
 QY 1789 AAGAAGATGTCCTGAAACCGAATCAAGTGGCCCATGTCAAGGCTGAGAGGAGCATCTG 1848
 DB 2487 AAGAAGATGTCCTGAAACCGAATCAAGTGGCCCATGTCAAGGCTGAGAGGAGCATCTG 2546
 QY 1849 GCTGAAGCAGACAAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGGACAAAGCAGC 1908
 DB 2547 GCCGAGCAGACAAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGGACAAAGCAGC 2606
 QY 1909 CTGTACTTTGTGAGTACATACATACAGGCGGGGATGATGAGCCTGCTGATCAGATG 1968
 DB 2607 CTGTACTTTGTGAGTACATACATACAGGCGGGGATGATGAGCCTGCTGATCAGATG 2666
 QY 1969 GAGTCTTCTCTGAGCAGCTGGCGGCTTCTACATTCAGAGTTGACCTGGCCATTGAA 2028
 DB 2667 GAGTCTTCTCTGAGCAGCTGGCGGCTTCTACATTCAGAGTTGACCTGGCCATTGAG 2726
 QY 2029 AGTGTCCACAGATGGGCTTTATCCACCGGACATCAAGCCTGACACATACTACTACAC 2088
 DB 2727 AGTGTCCACAGATGGGCTTATCCACCGGACATCAAGCCTGATACATTTTATAGAT 2786
 QY 2089 CTGGATGGTCAATATTAGCTGACAGATTTTGGCTCTGCACTGGATTCAGGTGGACTCAC 2148
 DB 2787 CTGGATGGTCAATATTAGCTGACAGATTTTGGCTCTGCACTGGATTCAGGTGGACTCAC 2846
 QY 2149 AATTCCAAGTACTACCAAGAAAGGAAACACATGAGACAGGACAGCATGGAGCCCGGTGAC 2208
 DB 2847 AATTCCAAGTACTACCAAGAAAGGAAACACATGAGACAGGACAGCATGGAGCCCGGTGAC 2906
 QY 2209 CTCTGGAGCAGATTTTCCAACTGTGCTGTGGAGACAGGTAAAGACCCCTGGAGCAGAGG 2268
 DB 2907 CTCTGGAGCAGATTTTCCAACTGTGCTGTGGAGACAGGTAAAGACCCCTGGAGCAGAGG 2966
 QY 2269 GCGCAGAGCAGACAGAGGTGCTTGGCAGATTTTGGCTGCTGGGACACCAATATACATC 2328
 DB 2967 GCGCAGAGCAGACAGAGGTGCTTGGCAGATTTTGGCTGCTGGGACCTCCAAATATACATC 3026
 QY 2329 GCTCCGAGGTGCTTCTCCCAAGAGGTACAGCAGCTCTGTGACTGCTGTGAGGCTCGGT 2388
 DB 3027 GCACCCGAGGTGCTTCTCCCAAGAGGTACACTCACTCTGTGACTGCTGTGAGGCTGTTGA 3086
 QY 2389 GTGATTTCTTTGAGATGCTGGTGGCAGCGCGCTTTCTTGGCCCCCCCCACACAGAG 2448

||||| 3087 GTGATCTCTTCAGATGCTGTGGGAGCGCCCTTTTGGCACCTACTCCACAGAA 3146
 Db
 QY 2449 ACAGAGCTGAAGTATCACTGAGAGCAGCGCTGCATATCCCTACGAGTGGCTC 2508
 Db 3147 ACCAGCTGAAGTATCACTGAGAGCAGCGCTGCATATCCCTACGAGTGGCTC 3206
 QY 2509 AGCGTGAAGCCGAGAGCTCATACAGAGCTGTGCTGCGCGCTGACTGCGCGCTGGGC 2568
 Db 3207 ACCCTGAGGCGAGGAGCTCATACAGAGCTGTGCTGCGCGAGACCGCTGGGG 3266
 QY 2569 AGGATGAGGCGAGAGCTCATACAGAGCAGCGCTGTCTTCAACACCATCGACTTTTCCCGT 2628
 Db 3267 CGGAATGAGGCGAGGAGCTCATACAGAGCAGCGCTGTCTTCAACACCATCGACTTTTCCCGT 3326
 QY 2629 GACATCCGAAGCAGGCTGACCTCATACAGAGCAGCGCTGTCTTCAACACCATCGACTTTTCCCGT 2688
 Db 3327 GACATCCGAAGCAGGCTGACCTCATACAGAGCAGCGCTGTCTTCAACACCATCGACTTTTCCCGT 3386
 QY 2689 AATTTTGACCCCGTGGATGAAGAAAGCCCTTGGCAGAGGCGGAGAGCGCCCAAG 2748
 Db 3387 AATTTGACCCCGTGGATGAAGAAAGCCCTTGGCAGAGGCGGAGAGCGCCCAAG 3446
 QY 2749 GCCTGGGACAGCTGGCTTCCCGGAGCAGAGCAGCGCTGTCTTCAACACCATCGACTTTTCCCGT 2808
 Db 3447 GCCTGGGACAGCTGGCTTCCCGGAGCAGAGCAGCGCTGTCTTCAACACCATCGACTTTTCCCGT 3506
 QY 2809 ACCTTCCGAGGCTTCTCGATGACAGAGCAGCGCTGTCTTCCCGTGGCGGAGCGCTCAG-- 2866
 Db 3507 ACCTTCCGAGGCTTCTCGATGACAGAGCAGCGCTGTCTTCCCGTGGCGGAGCGCTCAG-- 3566
 QY 2867 -----AGCCCGCAGAGAGTGCAGACCCAGGAGTGGGAGTGGAGGTCGGGCTG 2916
 Db 3567 GCAGAGGCTTCACAGCTGAGAGCTCAGATTTAGAAAGCTGTGATCTGGTGTGAGT 3626
 QY 2917 GAGGCTGCGAGCGCGGTGTACGTGTA 2942
 Db 3627 GAAGGCTGCGAGCGCTGTGTACGTGTA 3652

RESULT 6

AAI171765
 ID AAI171765 standard; DNA; 3533 BP.

XX AC AAI171765;

XX DT 18-JAN-2002 (first entry)

XX DE Human cancer-inhibiting gene Lats2 coding sequence.

XX KW Human; cancer-inhibitor; Lats2; kinase; cytosolic; ds.

XX OS Homo sapiens.

XX PN JP2001231565-A.

XX PD 28-AUG-2001.

XX PF 18-FEB-2000; 2000JP-0041818.

XX PR 18-FEB-2000; 2000JP-0041818.

XX PA (IGAK-) IGAKU SEIBUTSUGAKU KENKYUSHO KK.

XX PA (NOJII) NOJIMA H.

XX DR WPI; 2001-650995/75.

XX DR P-PSDB; AAG78992.

XX PT Novel human Lats2 gene useful for diagnosing cancer

XX PS Claim 1; page 17-18; 34pp; Japanese.

XX CC The present sequence is the coding sequence for human cancer-inhibiting

CC gene Lats2. Lats2 is a kinase and can be used for gene diagnosis of
 CC cancers.
 XX
 SQ Sequence 3533 BP; 831 A; 1096 C; 1011 G; 595 T; 0 other;
 Query Match 54.4%; Score 1717.2; DB 22; Length 3533;
 Best Local Similarity 77.2%; Pred. No. 0;
 Matches 2382; Conservative 0; Mismatches 548; Indels 156; Gaps 18;

QY 1 ATGAGAGCACCACCCGAAAGTTGGACCTTATCAAAAGCTCTCAGGGAATCGATATTC 60
 Db 190 ATGAGAGCACCACCCGAAAGTTGGACCTTATCAAAAGCTCTCAGGGAATCGATATTC 249
 QY 61 CTCCTGCTTTTGGCAACGAGTTCAGGAGCTTCGGAGCTGAGAGTGAACCGGAGATG 120
 Db 250 TTGTTGCTTTTGGCAACGAGTTCAGGAGCTTCGGAGCTGAGAGTGAACCGGAGATG 309
 QY 121 CTTGAGGAGTTGGTGAATGCGGCATGTGACAGGAGATGGTGGCAGAGCGCTCACGAG 180
 Db 310 CTGACAGGAATGTTGCAACGAGATGCGACAGGAGATGGTGGCAGAGCTCTCAAGCAG 369
 QY 181 ACGGCGAGTAGGATGATGAAAGCTGCTTGGAGTACATCAGTAAAGTGGCTACCTGAC 240
 Db 370 ACTGCGCAGCAGGAGCATCGAGCGCCCTGGAGTACATCAGCAAGTGGCTACCTGGAC 429
 QY 241 CCGAGGAATGAGCAGATGTTGCGAGTCAATCAAGCAGACCTCCCGAGGAAGGGCTCATG 300
 Db 430 CCGAGGAATGAGCAGATGTTGCGAGTCAATCAAGCAGACCTCCCGAGGAAGGGCTCATG 489
 QY 301 TCCACCCCGGTGACTCGGCGGCCAGTTTTCAGAGGCGCAGAGGAGAGCACTCCCATCTAC 360
 Db 490 CCAACCCCGGTGACTCGGCGGCCAGTTTTCAGAGGCGCAGAGGAGAGCACTCCCATCTAC 549
 QY 361 CACAGCTGGTGGTGGCAAACTACAG-----GGCCCCCGCCCA 399
 Db 550 CACAGCTGGTGGTGGCAAACTACAG-----GGCCCCCGCCCA 609
 QY 400 CTGGAGAGATGCGCGGCAATATTTAGACTTTCTCTTCTGAGCGCGGAGCGGCGACC 459
 Db 610 CTGGAGAGATGCGCGGCGCGCTGCTGAGTACCTTTTCCCGGAGTGGCGGCGGCGG 669
 QY 460 CACGTTGCCAGGCTCACAGCATCTCCCAAAGGTTACAGCAGCAGTACAGAGCAAGT 519
 Db 670 CCGGCGCCACAG---CACCAGCACCACCAAGGCTACGGTGCAGCGCTAGAGGACGCA 726
 QY 520 GCG-----CACTTCCGGGCGACACATATGCTGCTGCTCATCTACT-----ATCG 564
 Db 727 GGGGCACACTTCCCGCTGAGGCGCGCAGTACGCGGCGCGCAGCTGCTGCTGCTGGG 786
 QY 565 GAGCAGTCTGGGTATGGGTGCGAGCGCAGTTCTCTTCCAGAACAGCAGCAGCAGAT 624
 Db 787 GAACCCCTGGGCTACGAGTGCAGCGCGCGCTCTCTTCCAGAGCAAGAGCGCGCGGAG 846
 QY 625 GCC-----TATTCAGCATGGCC---AAGGCCAGGTTGGCGCTCC---CGCCAGCGCTC 672
 Db 847 ACCGGGGGTACGCGCAGCTGCGCAGAAAGGCGCAGGAGGAGCAGCGCGCGCGCTC 906
 QY 673 ACCTTTCTGCGCATGCTGGGTGTACTGCTGCGCAGCAGCAGCGCGCGCTTACCCCA 732
 Db 907 GCTTTCCACCCCTGCGCGCGGCTCTACGTTGCGCGCGCAGCAGCAGCAGCGCGCT 966
 QY 733 CTTGGGCGCCACCATTTACATGTTGGGCGCGCGGCTCCCGGCTTCTGCTGGCGAAGC 792
 Db 967 CCGTGGCGCCACCATGCTGCTGGGCTCCCGGCGCGGCTTCTGCTGGCGAAGCAGCAG 1026
 QY 793 TCTGCACAGGCTGTGCTGGCAGCGCTCCAGGAACAGCTCAATGCTGCTGCTGCTGCTG 852
 Db 1027 CCGCGCAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1086
 QY 853 GG-----CTCCAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 906
 Db 1087 AGCAGCAGCTCCGTCAGCAGTGGCGCGCTGCCACCTTGGCGCGCGGCTGCTGCTGCTG 1146

RESULT 7

AAZ61160
 ID AAZ61160 standard; cDNA; 1961 BP.
 AC AAZ61160;
 XX 30-MAY-2000 (first entry)
 DT CDNA S8171A encoding a partial protein kinase.
 XX Kinase activity; molecular weight marker; isoelectric focusing marker;
 KW peptide fragmentation control; cellular signal transduction; ss.
 XX
 OS Homo sapiens;
 XX
 FH Key Location/Qualifiers
 FT 3.1769
 FT /tag= a
 FT /note= "partial sequence"
 XX
 PN WO200008180-A2.
 XX
 PD 17-FEB-2000.
 XX
 PF 03-AUG-1999; 99WO-US17630.
 XX
 PR 04-AUG-1998; 98US-0095270.
 PR 11-SEP-1998; 98US-0099972.
 XX
 PA (IMV) IMMUNEX CORP.
 XX
 PI Virca GD, Bird TA, Anderson DM, Marken JS;
 XX
 DR WPI; 2000-195584/17.
 DR P-PSDB; AAY69162.
 XX
 PT New human kinase polypeptides and polynucleotides used as molecular
 PT weight markers and as controls for peptide fragmentation
 XX
 PS Claim 1; Page 9; 60pp; English.
 XX
 CC The present sequence encodes a partial polypeptide which has kinase
 CC activity. The kinase polynucleotides can be used to express the
 CC polypeptides, and as probes to identify nucleic acids encoding
 CC proteins having kinase activity. The kinase polypeptides and
 CC fragmented polypeptides are used as molecular weight and isoelectric
 CC focusing markers, and as controls for peptide fragmentation. They also
 CC have a number of therapeutic uses as kinases play a central role in
 CC cellular signal transduction. The polypeptides could also be used to
 CC identify binding partner proteins. The polypeptides can also be used as
 CC a reagent to identify any proteins that the polypeptide regulates, and
 CC proteins with which it might interact. The polypeptides may also be
 CC used for preparation of antibodies. The antibodies can be used in
 CC assays to detect the presence of the protein, and to purify the protein
 CC by immunoaffinity chromatography.
 XX
 SQ Sequence 1961 BP; 504 A; 550 C; 539 G; 368 T; 0 other;
 Query Match 41.0%; Score 1293; DB 21; Length 1961;
 Best Local Similarity 84.4%; Pred. No. 0;
 Matches 1497; Conservative 0; Mismatches 255; Indels 21; Gaps 3;
 QY 1185 CCCGCTGGATGTGGACTATGGCGCTCCGAGCGCAGGTGCCACCGCTCCGATATCCCAA 1244
 DB 2 CCCGCTGGAGTGGAGTACGAGGAGCCGAGCGAGGTGCCCGCTCCGCTCCCGGAA 61
 QY 1245 GCATCTGCTGCTGCCAGTAAGTCTGAGCAGTACAGCGTGGACCTGGAGCAGCTGTGCAC 1304
 DB 62 GCACCTGCTGCTGCCAGTACGAGCAGTAC-----GACCTGGACAGCTGTGCGC 115
 QY 1305 CAGTGTGACGACAGCTCTGAGAGGGG-----CACTGATCTAGAGGGGAGTGAACAGCCA 1361
 DB 116 AGGCATGGAGCAGAGCTCCGTCGGGGGCCCAACGAGCCCGGAGGCGGCAAGAGCGG 175

QY 1362 CAAAGTGCAGAGGAGACAAAGCTGGCAGAGACAAAAGCAGATTACAGCCTCCCGGT 1421
 DB 176 CAAAGCGCCAAAGGGGACAAAGAGGAAAGATATAAGAGCAGATTACAGCCTCCCGT 235
 QY 1422 GCCTGTCCGCAAGATAGCAGAGATGAGAGAGAGAGTCTCGATCAAGAGTTACTC 1481
 DB 236 TCCCGTCCGCAAAACAGCAGAGAGAGAGAGAGAGTCCAGCATCAAGAGTACTC 295
 QY 1482 CCCTTATGCTTCAATTTCTTCATGGAGCAACACAGCTGGAGAGATGTCATCAAAACCTACA 1541
 DB 296 GCATACGCTTTAAGTTCTTCATGGAGCAGCAGCTGGAGAGATGTCATCAAAACCTACA 355
 QY 1542 GCAGAGGTACGCGGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGA 1601
 DB 356 GCAGAGGTTAACCAGGAGGCTGAGCTGGAGCAAGAAATGGCCAAAGCTGGGCTCTGTGA 415
 QY 1602 GCGCAGCAGGAGCAGATGAGGAGATCTCTTACAGAGAGAGTCTAATCAACCGGCT 1661
 DB 416 AGCTGAGCAGGAGCAGATGCGGAGATCTCTTACCAAGAAAGTCTAATTACCAAGGTT 475
 QY 1662 GAAGAGGCCAAGATGGACAAGTCCATGTTCTGAAATCAAGACTCTAGGCATCGGTGC 1721
 DB 476 AAAGAGGCCAAGATGGACAAGTCTATGTTCTCAAGATCAAAACCTGGGATCGGTGC 535
 QY 1722 CTTTGGGAAAGTGTGCTTGTAAAGCTGGACACTACAGCTCTGTACGCCATGAAGAC 1781
 DB 536 CTTTGGAGAGTGTGCTTGTAAAGTGGACACTACAGCTCTGTACGCCATGAAGAC 595
 QY 1782 TCTCAGGAAGAAGATGCTCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGA 1841
 DB 596 CCTAAGGAAAAAGATGCTCTGAACCGGAATCAGGTGGCCCATGTCAAGGCTGAGAGGA 655
 QY 1842 CATCTGCTGAGCAGACAAATGAGTGGTGTGCTCAAACTCTACTACTCTCTCCAGACAA 1901
 DB 656 CATCTGCGCGAGGAGCAAAATGAGTGGTGTGCTCAAACTCTACTACTCTCTCCAGACAA 715
 QY 1902 GGACAGCTGTACTTTGTGATGGACTACATACAGCGGGGATATGATGAGCTGCTGAT 1961
 DB 716 AGACAGCTGTACTTTGTGATGGACTACATCTCTGTTGGGACATGATGAGCTGCTGAT 775
 QY 1962 CAGGATGAGGTCTTCCCTGAGCAGCTGGCCGCTTCTACATTTGAGAGTTGACCTGCG 2021
 DB 776 CCGGATGAGGTCTTCCCTGAGCAGCTGGCCGCTTCTACATTTGAGAGTTGACCTGCG 835
 QY 2022 CATTTAAAGTGTCCACAGATGGCTTTATCCACCGGACATCAAGCTTGAACAACTACT 2081
 DB 836 CATTTAGAGTGTCCACAGATGGCTTTATCCACCGGACATCAAGCTTGAACAACTACT 895
 QY 2082 CATCAGCTGGATGTCTATATTAAGCTGACAGATTTTGGCTTCTGCTGATTCAGGTG 2141
 DB 896 GATAGATCTGGATGTCTACATTAAGCTGACAGATTTTGGCTTCTGCTGATTCAGGTG 955
 QY 2142 GACTCAATTCAGTACTACCAAGAGGAAACACATGAGACAGGACAGATGAGCC 2201
 DB 956 GACTCAATTCAGTACTACCAAGAGGAAACACATGAGACAGGACAGATGAGCC 1015
 QY 2202 CGGTCACTCTGGGACAGTGTTCACACTGTCTGCTGTGGAGACAGTTAAACACCTGGA 2261
 DB 1016 CAGCCACCTCTGGGATGATGTCTTACTGCTGTGGGACAGGCTGAAGACCTAGA 1075
 QY 2262 GCAGAGGCGCAGAGCAGCAGAGGTGCTTGGCACAATCTTGTGGGACACCAAA 2321
 DB 1076 GCAGAGGCGCAGAGCAGCAGAGGTGCTTGGCACAATCTTGTGGGACACCAAA 1135
 QY 2322 TTACATCGCTCGGAGGTGCTTCCGGAAGGGTACAGGAGCTCTGTGACTGTGGAG 2381
 DB 1136 CTACATCGCACCAGGTGCTCTCCGGAAGGGTACACTCAACTCTGTGACTGTGGAG 1195
 QY 2382 CGTCTGTGTATCTCTTTGAGATGCTGTTGGGAGCGGCTTCTTGGCCCGCCACCC 2441
 DB 1196 TGTGTGAGTATCTCTTCGAGATGCTGTGGGAGCGGCGCTTTTGGACCTACTCC 1255
 QY 2442 CACAGAGAGCGAGCTGAAGGTGATCAACTGGGAGAGAGCGCTGCATATCCCTACGAGT 2501

Db 1256 CACAAACCCAGCTGAAGGTGATCAACTGGGAGAACACGCTCCACATTCAGGCCAGGT 1315
 Qy 2502 GAGGCTCAGCGCTGAGGCGGAGAGCTCATCAGCAAGCTGTGCTGGCGGCTGACTGCGG 2561
 Db 1316 GAAGCTGAGCCTGAGGCGGAGGACCTCATCAGCAAGCTGTGCTGGCGGAGACCG 1375
 Qy 2562 CTTGGGAGGATGGGCGGAGATGACCTCAAGGACACACCGCTTCTTCAACACCATCGACTT 2621
 Db 1376 CTTGGGCGGAATGGGCGGATGACCTGAAGGCCACACCGCTTCTTCAAGCGCATGACTT 1435
 Qy 2622 TTCCCGTCACATCCGAAGAGGCTGACCTAGTCCCGCACCATCAGCCACCCATGGA 2681
 Db 1436 CTCAGTGACATCCGGAAGCATCCAGCCCTTACCTTCCCGCACCATCAGCCACCCATGGA 1495
 Qy 2682 CACCTCCAAATTTGACCGGCTGGATGAAGAAAGCCCTGGCAGGCGGCGGAGAGAG 2741
 Db 1496 CACCTCGAATTCGACCCCGTAGATGAAGAAAGCCCTGGCAGGCGGCGGAGAGAG 1555
 Qy 2742 CGCAAGCGCTGGGACAGCTGGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTA 2801
 Db 1556 CACCAAGCGCTGGGACACACTCAGCTGCGCCCAATAAAGCATCTGAGCAGCGATTTA 1615
 Qy 2802 TGAGTTACCTTCGCGAGGCTTCGATGACACGCGCTATCCCTTCGCGTCCCGAGCC 2861
 Db 1616 CGAATTCACCTTCGCGAGGCTTCGATGACACATGGCTTACCCCTTCGATGCCCAAGCC 1675
 Qy 2862 CTCAG-----AGCCCGCAGAGGTGACAGCCAGCGGATCGGACTTGGGAAGG 2909
 Db 1676 TTCAGGAGCAGAAGCTTCACAGGCTGAGAGCTCAGATTTAGAAAGCTGTGATCTGGTGA 1735
 Qy 2910 TCGCGCGAGGCGTCCAGCGCGGTGACGTGA 2942
 Db 1736 TCAGACTGAAGGCTGCCAGCGCTGTGTAGTGTA 1768

RESULT 8

AAZ61158
 ID AAZ61158 standard; cDNA; 1498 BP.
 XX AC AAZ61158;
 XX 30-MAY-2000 (first entry)
 DE cDNA SS1771 encoding a human protein kinase.
 XX Kinase activity; molecular weight marker; isoelectric focusing marker;
 KW peptide fragmentation control; cellular signal transduction; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 2..1498
 FT /tag= a
 FT /product= "kinase"
 XX
 PN WO200008180-A2.
 XX 17-FEB-2000.
 XX
 XX 03-AUG-1999; 99WO-US17630.
 XX
 XX 04-AUG-1998; 98US-0095270.
 PR 11-SEP-1998; 98US-0099972.
 XX
 PA (IMMV) IMMUNEX CORP.
 XX
 PI Virca GD.; Bird TA, Anderson DM, Marken JS;
 XX
 DR WPI; 2000-195584/17.
 DR P-PSDB; AAY69160.
 XX
 PT New human kinase polypeptides and polynucleotides used as molecular

PT weight markers and as controls for peptide fragmentation
 XX Claim 1; Page 8-9; 60pp; English.
 XX The present sequence encodes a polypeptide which has kinase
 CC activity. The kinase polynucleotides can be used to express the
 CC polypeptides, and as probes to identify nucleic acids encoding
 CC proteins having kinase activity. The kinase polypeptides and
 CC fragmented polypeptides are used as molecular weight and isoelectric
 CC focusing markers, and as controls for peptide fragmentation. They also
 CC have a number of therapeutic uses as kinases play a central role in
 CC cellular signal transduction. The polypeptides could also be used to
 CC identify binding partner proteins. The polypeptides can also be used as
 CC a reagent to identify any proteins that the polypeptides regulates, and
 CC proteins with which it might interact. The polypeptides may also be
 CC used for preparation of antibodies. The antibodies can be used in
 CC assays to detect the presence of the protein, and to purify the protein
 CC by immunofluorescence chromatography.

XX SQ Sequence 1498 BP; 384 A; 417 C; 427 G; 270 T; 0 other;

Query Match 36.4%; Score 1148.4; DB 21; Length 1498;
 Best Local Similarity 86.3%; Pred. No. 3.5e-288;
 Matches 1295; Conservative 0; Mismatches 196; Indels 9; Gaps 2;

QY 1185 CCCGCTGGATGTGGACTATGCGGCTCCGAGCGGAGGTGCCACCGCTCCGTTATCCAAA 1244
 Db 4. CCCGCTGGAGTGGATGACGAGGCCAGACCGGAGGTGCCGCTCCGCCCTACCCGAA. 63
 QY 1245 GCATTTCTGCTGCCAGTAACTGTGAGCAGTACAGCGTGGACCTGGACAGCTGTGCAC 1304
 Db 64 GCACCTGCTGCTGCGCAGCAAGTCCGAGCAGTAC-----GACCTGGACAGCTGTGCGC 117
 QY 1305 CAGTGTGCAGCAGAGTCTGCGAGGGGG---CACTGTACTAGACGGGAGTGACAAGGCCA 1361
 Db 118 AGCATGGAGCAGAGCTCTGCGGGGCCCAAGCAGGCCCGGCGGCGGAGAGCCG 177
 QY 1362 CAAAGTGCAGAGGAGACAAAAGCTGGCAGAGACAAAAGCAGATTTCAGACCTCCCGGT 1421
 Db 178 CAAAGGCCCAAGGGGACAAAGGCGGAAAGGATAAAAGCAGATTTCAGACCTCTCCGT 237
 QY 1422 GCCTGCTCCGAAGATAGCAGAGTGAAGAGAGAGAGTCTGCGATCAAGATTACTC 1481
 Db 238 TCCGCTCCGCAAAACAGCAGAGAGAGAGAGAGTCAAGAGTCAAGAGTACTC 297
 QY 1482 CCCTATGCTTCAAAATCTTCATGGAGCAACAGCTGGAGAGTGTATCAAAAGCTTACCA 1541
 Db 298 GCATAGCCCTTTAAGTTCCTTCATGGAGCAGCAGCTGGAGAGTGTATCAAAAGCTTACCA 357
 QY 1542 GCAGAGGTGAGCGGAGGCTACAGCTGGAGAGAAATGCCAAAGCTGGGCTCTGTGA 1601
 Db 358 GCAGAGGTTTAAAGCGAGGCTGAGCTGGAGCAAGAAATGCCAAAGCTGGGCTCTGTGA 417
 QY 1602 GCGGAGCAGGAGCAGATGAGGAGAGTCTTACCAGAGAGTCTTAACACACCGGCT 1661
 Db 418 AGCTGAGCAGGAGCAGATGCGAAGATCTCTTACCAGAGAGTCTTAATTAACACAGGTT 477
 QY 1662 GAAGAGGCCCAAGATGACAAAGTCCATGTTGTGAAATCAAGACTCTAGGATCGGTGC 1721
 Db 478 AAGAGGGCCCAAGATGAGCAAGTCTATGTTGTCAAGATCAAAAGCTGGGATCGGTGC 537
 QY 1722 CTTTGGGAGAGTGTGCGCTTCGTTGTAAGCTGGACACTTCAGCTCTGTACGCCATGAAGAC 1781
 Db 538 CTTTGGGAGAGTGTGCGCTTCGTTGTAAGCTGGACACTTCAGCTCTGTACGCCATGAAGAC 597
 QY 1782 TCTCAGGAAGAAGATGCTCTGAACCGGATCAAGTGGGCCCATGTCAAGCTGAGAGGA 1841
 Db 598 CCTAAGGAAAAAGGATGTCTCTGAACCGGATCAAGTGGGCCCATGTCAAGCCGAGAGGA 657
 QY 1842 CATCTGGCTGAAGCAGACAAATGAGTGGGTGGTCAAACTCTACTACTCTCTCCAGGACAA 1901
 Db 658 CATCTGGCCGAGGAGCAGCAATGAGTGGGTGGTCAAACTCTACTACTCTCTCCAGGACAA 717

QY 1902 GGACAGCTGTACTTTGTGATGACTACATACCGGGGGGATATGATGAGCTGCTGAT 1961
 Db 718 AGACAGCTGTACTTTGTGATGACTACATACCGGGGGGATATGATGAGCTGCTGAT 777
 QY 1962 CAGGATGAGGTCCTCCCTGAGCAGCTGCGCGCTTCTACATTTGAGAGTTGACCTGGC 2021
 Db 778 CCGGATGAGGTCCTCCCTGAGCAGCTGCGCGCTTCTACATTTGAGAGTTGAGCTGCTG 837
 QY 2022 CATTGAAGAGTCCACAGATGGGCTTTATCCACCGGACATCAAGCTGACACATACT 2081
 Db 838 CATTGAGAGTCCACAGATGGGCTTTATCCACCGGACATCAAGCTGACACATTTT 897
 QY 2082 CATCCAGCTGGATGCTATTAAGCTGACAGATTTTGGCCTCTGCTGATTCAGGTG 2141
 Db 898 GATAGATCTGGATGCTACATTAAGCTGACAGATTTTGGCCTCTGCTGATTCAGGTG 957
 QY 2142 GACTCACAATTCAGTACTACAGAAAGGAGGACATGAGACAGGACATGGAGCC 2201
 Db 958 GACTCACAATTCAGTACTACAGAAAGGAGGACATGAGACAGGACATGGAGCC 1017
 QY 2202 CGGTCACTCTGGGACGATTTTCAACTGCTGCTGTGGACAGGTTAAAGACCTTGA 2261
 Db 1018 CAGCGACCTCTGGGATGATGCTTAACTGCTGCTGTGGGACAGGCTGAAGACCTTGA 1077
 QY 2262 GCAGAGGGCGGAGAGCAGCAGCAGGCTGCTGCGACATTTCTTGTGGGACACCAA 2321
 Db 1078 GCAGAGGGCGGAGAGCAGCAGCAGGCTGCTGCGACATTTCTTGTGGGACACCAA 1137
 QY 2322 TTACATGCTCGGAGGTCCTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAG 2381
 Db 1138 CTACATGCTCGGAGGTCCTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAG 1197
 QY 2382 CGTGGTGTGATTCCTTGTGATGCTGTTGGGACGCGCTTTCTTGGCCCCCACC 2441
 Db 1198 TGTGGGATGATTCCTTGTGATGCTGTTGGGACGCGCTTTCTTGGCCCCCACC 1257
 QY 2442 CACAGAGCGGAGTGAAGGTGATCAACTGGGAGAGCAGCTGCTGATATCCCTTACGAGT 2501
 Db 1258 CACAGAGCGGAGTGAAGGTGATCAACTGGGAGAGCAGCTGCTGATATCCCTTACGAGT 1317
 QY 2502 GAGGTCAGGCTGAGGCGGAGAGCTCATCAGAGTGTGCTGCGGCGGTGACTGCGG 2561
 Db 1318 GAACTGAGGCTGAGGCGGAGAGCTCATCAGAGTGTGCTGCGGCGGTGACTGCGG 1377
 QY 2562 CTTGGGAGGATGGGAGAGTCACTCAAGCAGCAGCTTCTTCAACACATGACTT 2621
 Db 1378 CTTGGGAGGATGGGAGAGTCACTCAAGCAGCAGCTTCTTCAACACATGACTT 1437
 QY 2622 TTCCGTCGATCCGAAAGCAGGCTGACCTTACCTCCACCATCAGCCACCCCATGGA 2681
 Db 1438 CTCCAGTCATCCGAAAGCAGGCTGACCTTACCTCCACCATCAGCCACCCCATGGA 1497

RESULT 9

ID ABK43457
 AC ABK43457 standard; cDNA: 1912 BP.

XX AC ABK43457;

XX DT 05-JUN-2002 (first entry)

XX DE DNA encoding novel central nervous system protein #37.

XX KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
 KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
 KW cardiac arrest; cerebrovascular disorder; ischaemia; angioneurosis;
 KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
 KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
 KW adenocarcinoma; reproductive system disorder; testicular feminisation;
 KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
 KW respiratory disorder; renal disorder; kidney failure; blood disorder;
 KW myocardial infarction; wound healing; cell proliferation; skin aging;
 KW food additive; food preservative; gene therapy; gene; ss.

XX OS Homo sapiens.
 XX PN WO200155318-A2.
 XX XX 02-AUG-2001.
 XX PF 17-JAN-2001; 2001WO-US01332.
 XX XX 31-JAN-2000; 2000US-0179065.
 PR 04-FEB-2000; 2000US-0180628.
 PR 24-FEB-2000; 2000US-0184664.
 PR 02-MAR-2000; 2000US-0186350.
 PR 16-MAR-2000; 2000US-0189874.
 PR 17-MAR-2000; 2000US-0190076.
 PR 18-APR-2000; 2000US-0198123.
 PR 19-MAY-2000; 2000US-0205515.
 PR 07-JUN-2000; 2000US-0209467.
 PR 28-JUN-2000; 2000US-0214886.
 PR 30-JUN-2000; 2000US-0215135.
 PR 07-JUL-2000; 2000US-0216647.
 PR 07-JUL-2000; 2000US-0216880.
 PR 11-JUL-2000; 2000US-0217487.
 PR 11-JUL-2000; 2000US-0217496.
 PR 14-JUL-2000; 2000US-0218290.
 PR 26-JUL-2000; 2000US-0220963.
 PR 26-JUL-2000; 2000US-0220964.
 PR 14-AUG-2000; 2000US-0224518.
 PR 14-AUG-2000; 2000US-0224519.
 PR 14-AUG-2000; 2000US-0225213.
 PR 14-AUG-2000; 2000US-0225214.
 PR 14-AUG-2000; 2000US-0225266.
 PR 14-AUG-2000; 2000US-0225267.
 PR 14-AUG-2000; 2000US-0225268.
 PR 14-AUG-2000; 2000US-0225270.
 PR 14-AUG-2000; 2000US-0225447.
 PR 14-AUG-2000; 2000US-0225757.
 PR 14-AUG-2000; 2000US-0225758.
 PR 14-AUG-2000; 2000US-0225759.
 PR 18-AUG-2000; 2000US-0226279.
 PR 22-AUG-2000; 2000US-0226681.
 PR 22-AUG-2000; 2000US-0226868.
 PR 23-AUG-2000; 2000US-0227182.
 PR 23-AUG-2000; 2000US-0227009.
 PR 30-AUG-2000; 2000US-0228924.
 PR 01-SEP-2000; 2000US-0229287.
 PR 01-SEP-2000; 2000US-0229343.
 PR 01-SEP-2000; 2000US-0229344.
 PR 01-SEP-2000; 2000US-0229345.
 PR 05-SEP-2000; 2000US-0229509.
 PR 05-SEP-2000; 2000US-0229513.
 PR 06-SEP-2000; 2000US-0230437.
 PR 06-SEP-2000; 2000US-0230438.
 PR 08-SEP-2000; 2000US-0231242.
 PR 08-SEP-2000; 2000US-0231243.
 PR 08-SEP-2000; 2000US-0231244.
 PR 08-SEP-2000; 2000US-0231413.
 PR 08-SEP-2000; 2000US-0231414.
 PR 08-SEP-2000; 2000US-0232080.
 PR 12-SEP-2000; 2000US-0232081.
 PR 12-SEP-2000; 2000US-0231968.
 PR 14-SEP-2000; 2000US-0232397.
 PR 14-SEP-2000; 2000US-0232398.
 PR 14-SEP-2000; 2000US-0232399.
 PR 14-SEP-2000; 2000US-0232400.
 PR 14-SEP-2000; 2000US-0232401.
 PR 14-SEP-2000; 2000US-0233063.
 PR 14-SEP-2000; 2000US-0233064.
 PR 14-SEP-2000; 2000US-0233065.
 PR 21-SEP-2000; 2000US-0234223.
 PR 21-SEP-2000; 2000US-0234274.
 PR 25-SEP-2000; 2000US-0234997.
 PR 25-SEP-2000; 2000US-0234998.

HUMAN - \ HUMAN GENOME SCI INC.

Barash SC; Ruben SM;

WPT: 2001-581633/65.

P-PSDB; AAU8/12/

New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives

claim 1. SEQ ID No 47: 837pp; English.

The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (III) encoded by (I), are used to treat a medical conditions and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angioneurosis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial infarction. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein, and

	Score	DB	Length
35.58;	1118.6;	23;	1912;

Query Match 85.98; Pred. No. 2.2e-280;

Best Local Similarity	0; Mismatches	194;	Indels	12;	Gaps
Matches 1257:	Conservative	0;	Mismatches	194;	Indels

TECHNICAL SERVICES 1237, CONSTRUCTION 1551

1492 TTCAAATTCCTTCATGGAGCAACACCGTGGAGAAATGTCATCAAAACCTACCAAGCAGAAAGGTC 15

TCTGTTTGGGTTCACATCGAGCCTCCAGGAGAAGGTT 60

1 TTTAAGTCTTCATGGAGCAGCACGTGGAGAA'TGTCATCAAAACCTACCAAGGAGCT

-----GCGCCCAATCCCCAAGCTGGGCTCTGTGAGGCCGAGCAG 161

1552 AGCCGGAGGCTACAGCTGGAGCAGGAATGCCCAAGCTCCGCTT

5' - TGGCGGCGCTCCGCAAGGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAG 120

.61 AACCGAGGCTGCAGCTGGAGCAAGGTTGGCGTCTT

1512 CAGCCGATCAGGAAGATCCTCTACCAGAAGGAGTCTAACTACAACCGGCTGAAGAGGGCC 167

1612 GAGCAGATGAGCGAATGGCTTCTTT
||||| ||||| ||||| ||||| |||||
180

127 GAGCAGATGCGGAAGATCCTCTACCGAAAGAGTCTAATTACACAGGTTAAAGAGGGCC, 180

TAT GAGCAGCTCCGCGTCTTTCTGGGAA 173

1672 AAGATGGACAAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGCC-TTGGGGA 175

TCTGGAGAA 240

181. AAGATGGACAAGTCTATGTTGTCAAGATCAAAACCCTGGGGATCGGTGCCCTTGGAGAT

...-CCGCGCCCTCTACGCCATGAAGACTCTCAGGAAG 179

1732 GTGTGCCCTCGCTTGTAGCTGGACACTTCATGACCCCAATGAGCCTCAT

[illegible]

241 GTGTGCCTTGCTTGTAAAGTGGACACTCACGCCCTGTACGCCGATG

185

1792 AAGGATGTCTGAACCGGATCCTGGCCGCTT
||||| ||||| ||||| ||||| ||||| |||||
||||| ||||| ||||| ||||| ||||| |||||

201 AAGGATGTCCTGAACCCGGAATCAGGTGGCCACGTCAGGCGGAGGACATCCTGGCC 360

301 AAGGATGTCCCTGAACCGCAGTGGTT

1852 GAAGCAGACAATGAGTGGTGGTCAAACTCTACTCTCTCCAGGACAAGGACAGCCIG 19

[illegible]

361 GAGGCAGACAA TGAGTGGTGCTACTCTACTCC TTTCCAAGACA AATGACACCG

...TCTGATCAGGATGGAG 19.

1912 TAC TTT GTG TGG ACTAC ATAC CAGG CCGG GATATG ATGAG CCGCTG TGGCTG T

[illegible]

421 TACTTTGTGATGGACTACATCCCTGGTGGGACATGATGAGGCTGCTT

10

Age Group	No opinion	Not a good idea	Good idea	Excellent idea
18-24	45%	15%	25%	15%
25-34	40%	20%	30%	10%
35-44	35%	25%	35%	5%
45-54	30%	30%	40%	0%
55-64	25%	35%	40%	0%
65-74	20%	40%	35%	5%
75+	15%	45%	30%	10%

QY 1972 GTCTCCCTGAGCACCTGCGCCCTTCTACATTGACAGAGTTGACCTGCGCATTTGAAGT 2031
 Db 481 GTCTCCCTGAGCACCTGCGCCCTTCTACATTGACAGAGTTGACCTGCGCATTTGAAGT 540
 QY 2032 GTCCACAAGATGGCTTTTATCCACCGGGACATCAAGCCTGACAACATCTCATCGACCTG 2091
 Db 541 GTCCACAAGATGGCTTTTATCCACCGGGACATCAAGCCTGACAACATCTCATCGACCTG 600
 QY 2092 GATGTCATATTAAGTCAGACATTTTGGCTCTGACCTGAGTTCAGTGGACTCAAT 2151
 Db 601 GATGTCATATTAAGTCAGACATTTTGGCTCTGACCTGAGTTCAGTGGACTCAAT 660
 QY 2152 TCAAGTACTACAGAAAGGAAACCATGAGACAGACAGATGAGAGCCCGGTGACCTC 2211
 Db 661 TCAAGTACTACAGAAAGGAAACCATGAGACAGACAGATGAGAGCCCGGTGACCTC 720
 QY 2212 TGGGAGATGTTTCCAACTGTCGCTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCG 2271
 Db 721 TGGGAGATGTTTCCAACTGTCGCTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCG 780
 QY 2272 CAGACAGCAGCAGAGGTCCTGGACACATTTCTTGTGGACACCAATTTACATCGCT 2331
 Db 781 CAGACAGCAGCAGAGGTCCTGGACACATTTCTTGTGGACACCAATTTACATCGCT 840
 QY 2332 CCGGAGGTGCTTCTCCGAAAGGTTACACGACGCTCTGACTGTTGGAGCGTGGTGTG 2391
 Db 841 CCGGAGGTGCTTCTCCGAAAGGTTACACGACGCTCTGACTGTTGGAGCGTGGTGTG 900
 QY 2392 ATTCTCTTTGAGATGCTGTTGGGAGCGCCCTTTCTTGGCCCGCCACCCACAGAGAG 2451
 Db 901 ATTCTCTTTGAGATGCTGTTGGGAGCGCCCTTTCTTGGCCCGCCACCCACAGAGAG 960
 QY 2452 CAGCTGAAGTGATCAACTGAGAGAGCAGCTGTCATATCCCTACGAGTGGGCTCAG 2511
 Db 961 CAGCTGAAGTGATCAACTGAGAGAGCAGCTGTCATATCCCTACGAGTGGGCTCAG 1020
 QY 2512 GCTGAGGCGCGAGACCTCATCAAGAGCTGTGCTGCGCGCTGACTGCGCGCTGGGCGG 2571
 Db 1021 GCTGAGGCGCGAGACCTCATCAAGAGCTGTGCTGCGCGCTGACTGCGCGCTGGGCGG 1080
 QY 2572 GATGGGCGAGATGACCTCAAGGACACACCGGTTCTTCAACACCATCGACTTTTCCCGT 2631
 Db 1081 GATGGGCGAGATGACCTCAAGGACACACCGGTTCTTCAACACCATCGACTTTTCCCGT 1140
 QY 2632 ATCCGAAGCAGCTGACCTTACCTCCACCATCAGCAGCCACCCATGACACCTCCAT 2691
 Db 1141 ATCCGAAGCAGCTGACCTTACCTCCACCATCAGCAGCCACCCATGACACCTCCAT 1200
 QY 2692 TTTGACCCCGTGGATGAAGAAAGCCCTTGGCAGAGGCGGAGAGAGCGCCAGGCG 2751
 Db 1201 TTTGACCCCGTGGATGAAGAAAGCCCTTGGCAGAGGCGGAGAGAGCGCCAGGCG 1260
 QY 2752 TGGGACACCTGCTTCCCGGAGCAGACAGATCCAGACAGCGCTTCTATGAGTTTAC 2811
 Db 1261 TGGGACACCTGCTTCCCGGAGCAGACAGATCCAGACAGCGCTTCTATGAGTTTAC 1320
 QY 2812 TTTCCGAGGTTCTGATCAACAGGCTATCCCTTCCCGTGGCCGAGCCCTCAG - - - - 2866
 Db 1321 TTTCCGAGGTTCTGATCAACAGGCTATCCCTTCCCGTGGCCGAGCCCTCAG - - - - 1380
 QY 2867 - - - - - AGCCCGGAGAGAGTGCAGACCCAGGGGATGCGGACTTGAAGGTCGCGCCGAG 2919
 Db 1381 GAAGCTTACAGGCTGAGAGCTCAGATTTTGAAGGCTCTGATCTGGTGCATCAGACTGAA 1440
 QY 2920 GCTGCGCAGCGGTGACTGTGA 2942
 Db 1441 GCTGCGCAGCGGTGACTGTGA 1463

RESULT 10
 AAD03989
 ID AAD03989 standard; cDNA; 2043 BP.
 XX

AC AAD03989;
 XX 02-JUL-2001 (first entry)
 DE Human protein tyrosine kinase receptor (PTK) cDNA from clone HMUBT71.
 KW Human: protein tyrosine kinase receptor; PTK; antibacterial; antiviral;
 KW antifungal; antiparasitic; sarcoidosis; inflammation; asthma; arthritis;
 KW cardiovascular disorder; severe combined immunodeficiency; SCID; vaccine;
 KW hyperproliferative disorder; atherosclerosis; brain disorder; leukaemia;
 KW multiple sclerosis; Alzheimer's disease; vascular dementia; thrombosis;
 KW neurodegenerative disorder; gastrointestinal disorder; cancer; lymphoma;
 KW blood coagulation disorder; trauma; cerebrovascular disorder; tendonitis;
 KW gene therapy; nephrotic syndrome; glomerulonephritis; allergy; neoplasm;
 KW musculo-skeletal disorder; Parkinson's disease; autoimmune disorder;
 KW behavioural disorder; renal disorder; ss.
 XX Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 248..1585
 FT /*tag= a
 FT /product= "Human PTK protein from clone HMUBT71"
 XX
 PN WO200129564-A1.
 PD 26-APR-2001.
 XX
 XX 12-OCT-2000; 2000WO-US28066.
 PF 15-OCT-1999; 99US-0159542.
 PR 17-NOV-1999; 99US-0165914.
 PR 14-MAR-2000; 2000US-0189027.
 XX
 PA (HUMA-) HUMAN GENOME SCI INC.
 XX
 PI Ruben SM, Shi Y, Young PE, Ni J;
 XX
 DR WPI: 2001-290976/30.
 DR P-PSDB; AAE00663.
 XX
 PT New isolated nucleic acid molecule for producing human protein tyrosine
 PT kinase receptor
 XX
 PS Claim 1; Page 260-261; 288pp; English.
 CC The present cDNA sequence encodes human protein tyrosine kinase receptor
 CC (PTK) from clone HMUBT71.
 CC The invention relates to human protein tyrosine kinase (PTK) receptors
 CC and their corresponding cDNA molecules. PTK receptors are useful in
 CC providing immunological probes for differential identification of the
 CC tissues or cell types present in a biological sample. PTK is used in
 CC methods for the diagnosis, prevention and treatment of various disorders
 CC related to PTK such as immune system disorders (severe combined
 CC immunodeficiency (SCID), inflammation); hyperproliferative disorders
 CC (neoplasm, sarcoidosis); cardiovascular disorders (arrhythmia,
 CC atherosclerosis); central nervous system disorders (multiple sclerosis),
 CC neurodegenerative disorders (Alzheimer's disease, Parkinson's disease);
 CC blood coagulation disorders (thrombocytopenia); autoimmune disorders
 CC (biliary cirrhosis, Crohn's disease); respiratory disorders (asthma,
 CC allergy); gastrointestinal disorders (inflammatory bowel disease);
 CC cerebrovascular disorders (thrombosis, vascular dementia); brain
 CC disorders (phenylketonuria); cancers (such as ovarian, lung, bladder,
 CC liver, breast and lymphomas); behavioural disorders (Tourette's
 CC syndrome); musculo-skeletal disorders (arthritis, trauma, tendonitis);
 CC renal disorders (nephrotic syndrome, glomerulonephritis); metastases of
 CC malignancies and related disorders (leukaemia, multiple myeloma); and
 CC infections caused by bacteria, viruses, fungi and parasites. PTK is also
 CC useful for screening therapeutic compounds. PTK is used as an antigen in
 CC a vaccine to raise an immune response against infectious disease. PTK
 CC nucleic acids are useful in gene therapy.
 XX
 XX Sequence 2043 BP; 551 A; 528 C; 495 G; 469 T; 0 other;

QY	2605	TTCAACACCATCGACTTTTCCGCTGACATCCGAAGCAGGCTGCACCCCTAGCTGCCCAACC	2664		
Db	1235	TTTCCGCGCATGACTTCTCCAGTGACATCCGGAAGCAGCAGCCCTCTAGTTCCCAACC	1294		
QY	2665	ATCAGCCACCCCATGGACACCTCCAAATTTTACCCGGTGGATGAAGAAAGCCCTGGCAC	2724		
Db	1295	ATCAGCCACCCCATGGACACCTCCAAATTTTACCCGGTGGATGAAGAAAGCCCTGGAAC	1354		
QY	2725	GAGGCCAGCGGAGAGCGCCCAAGCCCTGGGACACGCTGGCCCTCCGCCAGCAGCAAGCAT	2784		
Db	1355	GATGCCAGCGAGGTAGCACCAGCCCTGGGACACACTCACCCTGGCCCAATAACAAGCAT	1414		
QY	2785	CCAGAGCAGCGCTTCTTATGAGTTTCACTTCCGAGGTTCTTCGATGACACGCGCTATCCC	2844		
Db	1415	CCTGAGCAGCATTTTACGATTCACCTTCGGAAGTTCTTTGATGACAATGGCTACCCC	1474		
QY	2845	TTCCGGTGGCCGAAGCCCTCAG-----AGCCCGCAGAGAGTGCAGACCCAGGG	2892		
Db	1475	TTTCGATGCCCAAGCCCTTCAGGAGCAGAAGCTTCACAGCTGAGAGCTCAGATTAGAA	1534		
QY	2893	GATCGCGACTTGGAAAGGTGGCGCGGAGGCTGCCAGCCCGTGTACTGTGA	2942		
Db	1535	AGCTCTGATCTGTGTGATCAGATGAAGCTGCCAGCTGTGTACTGTGA	1584		
RESULT 11					
AAZ51506					
ID	AAZ51506 standard; DNA; 3213 BP.				
AC	AAZ51506;				
XX					
DT	21-JUN-2000 (first entry)				
XX					
DE	Mouse Lats (large tumour suppressor) DNA.				
XX					
KW	Mouse; Lats; large tumour suppressor; cytostatic; vulnery;				
KW	cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;				
KW	treatment; prevention; screening; cancer; skin; ovarian tumour;				
KW	soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;				
KW	LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;				
KW	dysplasia; degenerative disorder; growth deficiency; physical trauma;				
KW	hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.				
XX					
OS	Mus musculus.				
XX					
FH	Key	Location/Qualifiers			
FT	CDS	1..2889			
FT		/*tag= a			
FT		/product= "Lats protein"			
FT		/partial			
XX					
PN	WO200010602-A1.				
XX					
PD	02-MAR-2000.				
XX					
PF	18-AUG-1999; 99WO-US19068.				
XX					
PR	18-AUG-1998; 98US-0096996.				
PR	18-AUG-1998; 98US-0096997.				
XX					
XX					
PA	(UYUA) UNIV YALE.				
XX					
PI	Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;				
PI	Turenchalk GS, Stewart RA;				
XX					
XX					
DR	WPI: 2000-246496/21.				
XX	P-PSDB; AAY70391.				
XX					
PT	Use of lats proteins, complexes of lats and cdc2 for treating cancer				
PT	that is refractory to treatment by standard chemotherapy and radiation				
PT	therapy, and disorders associated with aberrant levels of cdc2 activity				
XX					

Claim 44; Fig 13; 134pp; English.

The present sequence is a DNA encoding mouse Lats (large tumour suppressor) protein which is a cell overproliferation inhibitor and a negative regulator of cell cycle-dependent kinase cdc2/cyclin A. The present sequence is useful for treating cancer that is refractory to standard chemotherapy or radiation therapy such as hyperplasia, metaplasia, or dysplasia, and disorders associated with aberrant levels of cdc2 activity. Conditions treated by promoting cdc2 function include degenerative disorders, growth deficiencies, hypoproliferative disorders, physical trauma, lesions, and wounds. An animal model preferably a mouse, in which a lats gene has been disrupted by homologous recombination, e.g. a lats knock-out mouse, is used for screening compounds that can be used to treat or prevent cancer, particularly skin cancer, soft tissue sarcomas and ovarian tumours, and disorders associated with pituitary dysfunction e.g. luteinizing hormone (LH) hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

Sequence 3213 BP; 946 A; 764 C; 696 G; 807 T; 0 other;

Query Match 25.8%; Score 812.6; DB 21; Length 3213;
Best Local Similarity 65.1%; Pred. No. 1.1e-200;
Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

Qy 985 CAACCTGAGCCCTCACTGCCCGCCCCCAACACGGTCAACCGCGTACCGCCGCGCACATC 1044
Db 919 CACTCTGTAATCTTCAGCCTTCTGCCACTACAGTCACTGCATCACACCGCGTCTTAT 978
Qy 1045 CTTACACCTGTGAAGAGCGTGGTGTGCTGGCGCCGAGCCCCAGACAGCGCGTGGGGCC 1104
Db 979 CAACAGCCCGGTGAAGAGCATCGCGTCTGAAACAGAGCTGCAGACTGCTTTAGCCCCA 1038
Qy 1105 TCGACACCCCGCTGGTGGCTGCGCCACAGCACCTGCCACTGAGAGCGCTGGAGCAGAG 1164
Db 1039 ACCATCCTCTTGTGATGCCACAGCCAGTTTCAGACTGTTACAGCTACCCCTTTTCTGAG 1098
Qy 1165 GAGGCGACGCGAGCGCCACACCCCGCTGGATGAGTGGGCTGCGGCTCCGAGCGCAGGTGC 1224
Db 1099 GGTACAGCTTCAAGTGTCCCTGTATCCACCTGTTGCTGAAGCTTCCAAAGCTATCAAGT 1158
Qy 1225 CCACCGCTCGTATCAAGACACTTGTCTGCTGCCAGTAGTCTGACGAGTACAGCGTG 1284
Db 1159 CCACCGCCGCTTATCCAAACATCTGCTACACCAAAACCCATCTGCTCCCTCCATATAG 1218
Qy 1285 GACCTGGACAGCTGTGCACCACTGTGCAGCAGAGTGTGCGAGGGGCGCTAGCTAGAC 1344
Db 1219 TCAGTAAGTAAAGCTGCAAGATGAACAGCTAGCTTACCAAGGAAGATGATAGTAG 1278
Qy 1345 GGGAGTGACAGGCCACAAAGGTGCGAAGGAGACAAAGCTGGCAGACACAAAGCAG 1404
Db 1279 AAGAGTGGGACAGT-----GCTGACTCTGGGGATAAAGAAAGAAACAG 1323
Qy 1405 ATTCAGACTCCCGGTCCTGTCGGAAGATAGCAGAGATGAAGACAGAGAGAGTCT 1464
Db 1324 ATTAACACTTCCACTATCACTGTTCCGAAACAAAGAAAGATGAAGAACAGAGAGTCT 1383
Qy 1465 CGCATCAAGAGTACTCCCTTATGCCCTTCAATTTCTTATGGAGCAACAGTGGAGAT 1524
Db 1384 CGGATTAGAGTACTCCACAGCGCTTAAAGTCTTATCGAGCAGCAGCAGTAGAGAAC 1443
Qy 1525 GTCATCAAAACCTACAGCAGAGGCTACGCGGAGGCTACAGCTGGAGCAGGAATGGCC 1584
Db 1444 GTCCTGAAGTCTCATCAGCAGCGTGTGCATCGGAAGAAGCAGCTAGAAATGAATGATG 1503
Qy 1585 AAGCTGGGCTCTGTGAGCGCCAGGAGGAGCAGATGAGGAGAGATCTCTACCAAGAGAG 1644
Db 1504 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAG 1563
Qy 1645 TCTAACTACAACCGGCTCAAGAGGCGCAAGATGGACAGTCCATGTTTGTCAAAATCAAG 1704
Db 1564 TCTAACTATATCGTCTTAAAGGGCTTAAATGGACAGTCTATGTTTGTAAAGATAAG 1623
Qy 1705 ACTTAGGCATCGTGGCTTTGGGAGAGTGTGCCTGCTGTAAGCTGGACACTCACGCT 1764

Db 1624 ACATTAGGAATAGGAGCGTTTGGTGAAGTCTGTAGCAAGAAAGTGCATACAAAGCT 1683
Qy 1765 CTGTACGCCATGAGACTCTCAGGAAGAGAGATGCTTGAACCGAATCAAGTGGCCCAT 1824
Db 1684 TTGTATGCAACAAGAGACTCTTCGAAAGAAAGACGTTCTGCCGAAATCAGTGGCTCAT 1743
Qy 1825 GTCAGGCTGAGAGGAGCATCTGGCTGAAGCAGACAAATGAGTGGTGGTCAAACTCTAC 1884
Db 1744 GTGAAGCGGAGAGGGATATCTTAGCAGAACCCGACATGAGTGGTGGTCCGCTGTAC 1803
Qy 1885 TACTCTTCCAGGAAAGAGCAGCCTGTACTTGTGTGAGTACTACATACCAGCGGGAT 1944
Db 1804 TACTCTTCCAGGAAAGAGCAGCCTGTACTTGTGTGAGTACTACATTCCTCGGGGAT 1863
Qy 1945 ATGATGACCTGCTGATCAGATGAGTGGTCTTCCCTGAGCAGCTGGCCCGTCTTACAT 2004
Db 1864 ATGATGACCTTATTAATAGATGGCATCTTCTGAAATCTGGCAGATTTACATA 1923
Qy 2005 GCAGATGACCTTGGCCCTGAAAGTGTCCACAGATGGCTTATCCACCGGACATC 2064
Db 1924 GCAGACTTACCTGTGACAGTTGAAAGTGTCTAATAATGGTCTTATTCATAGATAT 1983
Qy 2065 AAGCTGACAACTACTCATCATCGACTGGATGGTGTATTAAGCTGACAGATTTTGGCCTC 2124
Db 1984 AAACCTGATAACATTTTGTATGACCGTGTGCCATATTAATTAATGACTGACTTTGGCTG 2043
Qy 2125 TGCATGATTCAGGTGAGCTCAATTTCCAAAGTACTACCAAGAAAGGAAACACATGAGA 2184
Db 2044 TGCATGCTTTCAGATGACACATGACTCCAAAGTACTACCAAGTGGGGATGCCACGG 2103
Qy 2185 CAGGACAGCATGGAGCCCGGTGACCTCTGGGACGATGTTTCCAACTGCTGCTGGAGAC 2244
Db 2104 CAAGATGACATGGATTTTCAGTAAAGATGGGAGATCTTCCATTTGCTGGTGGGAC 2163
Qy 2245 AGGTTAAAGACCTTGGAGCAGAGCGGCGAGACAGCAGCAGAGAGTGTGCTGGCAATCT 2304
Db 2164 AGACTGAAGCCACTGGAGCGGAGAGTGTCTGCCAGCAGCAGCAGTGTCTAGCCATCT 2223
Qy 2305 CTTGTCGGGACACCAATTTACATGCTCCCGAGGTGCTTCCGCAAGAGGTACACGAG 2364
Db 2224 CTGTTGGAGTCCCAATTTATTTGACCTGAAGTGTCTGCGAAGAGATATACACAG 2283
Qy 2365 CTCTGTGACTGGTGGAGCTGGTGTGATTTCTTTGAGATGCTGTTGGGAGCGGCT 2424
Db 2284 CTGTTGACTGGTGGAGTGTGTTGTTATTTCTTTGAAATGTTGGTGGGACACCTCT 2343
Qy 2425 TTTTGGCCCCCCCCACACAGCAGCAGCAGTGAAGTGTGATCAACTGGGAGCAGCAGCTG 2484
Db 2344 TTTTGGCACAACCCCAATTTAGAAACACAAATGAAGTTATCATCTGGCAAACTTCTCTA 2403
Qy 2485 CATATCCCTACGAGGTGAGCTCAGCGCTGAGCGCCGAGACCTCATCAGAGCTGTGC 2544
Db 2404 CATATCCCTCCTCAAGCTTAAGCTGAGTCTGAAGCTCTGACCTCATTTATCAAACTGT 2463
Qy 2545 TCGCGGCTGACTGCGGCTGGGAGGATGGGAGATGAGCTCAAGGACACCCGCTTC 2604
Db 2464 CGAGGACCAAGAGCCGCTCGGCAAGAACCGTGTGATGAGATGAAGGCTCATCATTT 2523
Qy 2605 TTCAACACCATCTGACTTTTCCCGTGACATCCGAAAGAGGCTGCACCTCTAGTCCCCACC 2664
Db 2524 TTTAAGACCATCGATTTCTCTAGTGTCTGAGACAGCAGTGTGCTTATCATATCCCTAAA 2583
Qy 2665 ATCAGCCACCCATGGACACCTCCAAATTTTGAACCGGTGGATGAAGAAAGCCCTGGCAG 2724
Db 2584 ATCAGCATCCCAACAGATACATCAATTTGACCTCTGATCTGTGATTAATTTGGAGC 2643
Qy 2725 GAGCCAGCGAGAG---AGCGCCAAAGCGCTGGGACAGCTGGCTCCCCCAGCAGCAAG 2781
Db 2644 GATGGCAGGAGGAGGAAATATCAGTGACACTCTGACCGGATGTTATTAATTTGGAG 2703
Qy 2782 CATCCAGACGCGCTTCTATGAGTTCACCTTCCGCGAGTCTTTCGATGACAACGGCTAT 2841

Fri Jan 17 11:17:03 2003

Db 2704 CACCCGAGCAGCCTTCTATGAGTTCACCTTTCGGAGGTTTTTGTGATGACAATGGCTAC 2763
 QY 2842 CCCTTCGGTGGCCGAGCCTCAGAGCCGCGAGAGT 2880
 Db 2764 CCATATATATCAAGCCTATTGAGTATGATACATT 2802

RESULT 12
 AAT42119
 ID AAT42119 standard; cDNA; 3213 BP.
 XX
 AC AAT42119;
 XX
 DT 31-JAN-1997 (first entry)
 XX
 DE M-lats gene encoding large tumour suppressor.
 XX
 KW Mouse; m-lats gene; large tumour suppressor; fetal brain;
 KW protein-serine/threonine-kinase; cell proliferation; antisense;
 KW dominant-negative; cancer; degenerative disorder; trauma;
 KW growth deficiency; therapy; antitumour; vulnerability; diagnostic;
 KW transgenic plant; transgenic animal; growth; senescence; ds.
 XX
 OS Mus musculus.
 XX
 FH Key Location/Qualifiers
 CDS 1..2889
 FT /*tag- a
 FT /product= m-lats protein
 XX
 PN W09630402-A1.
 XX
 PD 03-OCT-1996.
 XX
 PF 26-MAR-1996; 96MO-US04101.
 XX
 PR 27-MAR-1995; 95US-041111.
 XX
 PA (UYUA) UNIV YALE.
 XX
 PI Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX
 DR WPI: 1996-455275/45.
 DR P-PSDB; AAW05179.
 XX
 XX New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX
 PS Disclosure; Page 126-130; 215pp; English.
 XX
 CC This sequence encodes a mouse large tumour suppressor m-lats protein,
 CC and has been isolated from a newborn mouse brain phage lambda-ZAP
 CC cDNA library using a 2.2-kb DNA probe from the Drosophila lats gene
 CC (AAT42117). A homologous mouse sequence has also been isolated
 CC (m-lats2, AAT42120). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX
 SQ Sequence 3213 BP; 946 A; 762 C; 697 G; 808 T; 0 other;

Query Match 25.78; Score 809.4; DB 17; Length 3213;
 Best Local Similarity 65.08; Pred. No. 7.6e-200;
 Matches 1235; Conservative 0; Mismatches 646; Indels 18; Gaps 2;
 QY 985 CAACTGAGCCCTACTCCCGCCGCCCAACACGCTCAGCGCGGCGGACATC 1044

919 CACTCTGCTAATCTCAGCCCTTCTGCCACTACAGTCACTGCGCATCACACCCGCTCCTATT 978
 1045 CTTACCCCTGTGAAGAGCGTGGTGTGCTGCGGCCGAGCCACAGACAGCCGTGGGGCCC 1104
 979 CAACAGCCCGTGAAGAGCATGCGCGCTCTGAACACAGAGCTGCAGACTGCTTTAGCCCA 1038
 1105 TCGCACCCCGCTGGTGGCTGCGCCACAGCAGCAGCTTCCAGCTGAGAGCTTGAGAGAG 1164
 1039 ACCATCTCTTGGATGCCACAGCCAGTTTACAGCTGTTTCCAGCTTCTTTCTGAG 1098
 1165 GAGGGCAGCCAGCCACACCCCGCTGGATGTGGAGTATGGCGGCTCCGAGCGCAGGTGC 1224
 1099 GGTACAGCTTCAAGTGTGCTGTATCCACCTGTTGCTGAAGCTCCAGCTATCAAGT 1158
 1225 CCACCGCTCGTATCCAAAGCAGCTTCTGCTGCCAGTAAGTCTGAGAGTACAGCGTG 1284
 1159 CCACCGCCCTTATCCAAACATCTGCTACACCAAAACCCATCTGTCTCCATATGAG 1218
 1285 GACCTGAGCAGCCTGTGCACAGTGTGCAGAGTGTGGAGGGGCGACTGATCTAGAC 1344
 1219 TCAGTAAGTAAGCCTGCAAGATGAACAGCTTAGCTTACCCAGGAAGATGATAGTG 1278
 1345 GGGAGTGACAAGAGCCACAAAGGTGCGAAGGAGACAAAGCTGCGAGAGACAAAAAGCAG 1404
 1279 AAGAGTGGGACAGT-----GGTACTCTGGGATAAAGAAAAAGAACAG 1323
 1405 ATTACAGACTCCCGGTGCTTCCGCAAGATAGCAGATGAGAGAGAGAGAGTCT 1464
 1324 ATTACAACCTTCACTTATCAGTGTTCGGAAGAAAGAGATGAAGAACAGAGAGTCT 1383
 1465 CGATCAAGAGTACTCCGCTTATGCTTCAATTTCTATGGAGCAACAGCTGGAGAAT 1524
 1384 CGGATTCAGAGTACTCCCGCAGCCCTTAAAGTCTTTCATGGAGCAGCAGTGAAGAC 1443
 1525 GTCATCAAAACCTTACCAGAGAGTCCGCGAGGCTACAGCTGGAGCAGAGAAATGCC 1584
 1444 GTCTGAAGTCTCATCAGCAGCGTCTGCTCGAAGAGAGAGCAGCTAGAAATGAATGATG 1503
 1585 AAGCTGGCTCTGTGAGCCGAGCAGCAGATGAGAGAGATCCTCTACCAGAGAGAG 1644
 1504 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAGAGATGCTTTGCCAGAGAG 1563
 1645 TCTAACTACAACCGCTGAAGAGGCGCAAGATGGAGAGTCCATGTTGTGAAGATCAAG 1704
 1564 TCTAACTATATTCGCTTAAAGGGCTAAATGGAGAGTCTATGTTGTAAGATGAAG 1623
 1705 ACTTAGGATCGTGGCTTGGGAGAGTGTGCTCGCTTGTAAAGCTGGAGACTCACCT 1764
 1624 ACATTAGGAATAGAGCGTTTGGTGAAGTCTGTCTAGCAAGAAAGTGCATATAAGCT 1683
 1765 CTGTACCCCATGAAGACTCTCAGGAAGAGAGTGTCTGAACCGGANTCAAGTGGCCAT 1824
 1584 TTGTATCAACAAAGACTTTCGAAAGAAAGAGCTTCTGCTCGAATCAGGTGGCTCAT 1743
 1825 GTCAAGCTGAGAGGAGCATCTTGGCTGAAGAGACAAATGAGTGGGTGCTCAAACTCTAC 1884
 1744 GTGAAGCGGAGGAGATATCTAGCAGAGAGCCGACAAATGAGTGGGTGGTCCGCTGTAC 1803
 1885 TACTCTTCCAGGACAAAGAGCAGCTCTACTTTGTGATGAGTACATACAGCGGGGAT 1944
 1804 TACTCTTCCAGGACAAAGAGCAACTTGTACTTTGTGATGAGTACATCTCTGGGGGAT 1863
 1945 ATGATGAGCTGTGATCAGGATGGAGTCTTCCCTCAGCAGCTGGCCGCTTCTACATT 2004
 1864 ATGATGAGCTTATTAATTAGAATGGGATCTTCTCTGAAATCTTGGCAGCATTTACATA 1923
 2005 GCAGAGTTGACCCCTGGCCATTGAAAGTGTCCAAAGATGGGCTTTATCCACCGGACATC 2064
 1924 GCAGAACTTACCTGTGCAAGTTGAAAGTGTTCATAAAATGGGTTTTTATTCATAGAGATT 1983
 2065 AAGCCTGACACATCTACTCATGAGCTGATGATATTAAGCTGACAGATTTTGGGCTC 2124
 1984 AAACCTGATAACATTTTATTGATGACCGTGTATGCCATATTAATTAAGTACTGACTTTGGCTG 2043

Qy	2125	TGCACTGGATTGAGTGGAGTCAACAATTCAGTACTACCAAGAAAGGAAACACATGAGA	2184
Db	2044	TGCACTGGCTTCAGATGGACACATCACTCAAGTACTACCAAGTGGGATCAACCCACGG	2103
Qy	2185	CAGGACACATGGAGCCCGGTGACCTCTGGGACGATGTTTCAACTGTGCGTGGAGAC	2244
Db	2104	CAAGATACATGGATTTTCAGTAAACAAATGGGAGATCTTCCAAATTTGCGTGGGAC	2163
Qy	2245	AGGTTAAAGACCTGGAGCAGAGGGCGAGAGCAGACAGAGTGTGCTGGCACAATCT	2304
Db	2164	AGACTGAAGCCACTGGAGCGGAGAGCTGCTCCAGCAGCAGCGATGTAGCCCAATCT	2223
Qy	2305	CTTGTGGGACACCAATATACATCGCTCCGGAGTGTCTTCGCAAAAGGTACACGAG	2364
Db	2224	CTGGTTGGGACTCCCAATATATATTCACCTGAAGTGTACTGCGAAGAGGATATACACAG	2283
Qy	2365	CTCTGTGACTGGTGGAGCGTGGTGTATCTCTTTGAGATGCTGGTGGGAGCGCT	2424
Db	2284	CTGTGTGACTGGTGGAGTGTGGTGTATCTTTGTGAATGTTGGTGGGACAACTCT	2343
Qy	2425	TTCTTGGCCCCACCCACACAGAGACGAGCTGAAGGTGATCAACTGGGAGAGCAGCTG	2484
Db	2344	TTCTTGGCACAACCCCAATTTAGAACACAATGAGGTATCATCTGGCAAACTTCTCTA	2403
Qy	2485	CATATCCCTAGCAGTGGAGTGGCTCAGCGCTGAGGCGCCGAGACCTCATCAGCAAGCTGTGC	2544
Db	2404	CACATCCCTCTCAAGCTAAGCTGAGTCTCTGAAGCTCTGACCTCATTTATCAAACTGTGT	2463
Qy	2545	TGCGGGGTGACTGCGCTGGGAGGAGGATGGGAGATGACCTCAAGGCACACCCGTTTC	2604
Db	2464	CGAGGACACAGAGACCGCTCGGCAAGACGCTGCTGATGAGATAAGGCTCATCAATT	2523
Qy	2605	TTCAACACCATCGATTTTCCGTGACATCCGAAAGCAGGCTGCACCTACGTCGCCACC	2664
Db	2524	TTTAAGACCATCGATTTCTTAGTGATCTGAGACAGCAGTCTGCTTATACATCCCTAAA	2583
Qy	2665	ATCAGCCACCCCATGGACACCTCCCAATTTTGACCGGTGGATGAAGAAAGCCCTGGCAC	2724
Db	2584	ATCAGGCATCCACAGATACATCCCAATTTTCGACCTGTGATCCTGATAAATTTGGAGC	2643
Qy	2725	GAGGCCAGCGGAGAG-AGGCCCAAGCCCTGGGACAGCTGGCCCTCCCGCAGCAGCAAG	2781
Db	2644	GATGCGCAGGAGGAGGAAATATCAGTGACACCTCTGAGCGGATGTTATATAAATGGGAG	2703
Qy	2782	CATCCAGAGCAGCTTCTATGAGTTACCTTCGCGAGTCTTCGATGACAAGCGCTAT	2841
Db	2704	CACCCGAGCAGCTTCTATGAGTTACGTTTCGAGGTTTTCGATGACAATGGCTAC	2763
Qy	2842	CCCTTCGGTCCCGAAGCCCTCAGAGCCCGCAGAGAGT	2880
Db	2764	CCATATAATTATCCAAAGCCTATTGAGTATGAATACATT	2802

RESULT 13

AAT42118

ID AAT42118 standard; cdna; 3984 BP.

XX AC AAT42118;

XX AC AAT42118;

XX DT 31-JAN-1997 (first entry)

XX DE H-lats gene encoding large tumour suppressor.

XX KW Human; h-lats gene; large tumour suppressor; fetal brain;

XX KW protein-serine/threonine-kinase; cell proliferation; antisense;

XX KW dominant-negative; cancer; degenerative disorder; trauma;

XX KW growth deficiency; therapy; antitumour; vulnerability; diagnostic;

XX KW transgenic plant; transgenic animal; growth; senescence; ds.

XX OS Homo sapiens.

XX FH

XX key

XX Location/Qualifiers

FT CDS

FT 231..3623

FT /*tag= a

FT /product= Lats protein

PN WO9630402-A1.

XX 03-OCT-1996.

XX 26-MAR-1996; 96WO-US04101.

XX 27-MAR-1995; 95US-0411111.

XX (UYUA) UNIV YALE.

XX Tao W, Wang W, Xu T, Yu W, Zhang S;

XX WPI; 1996-455275/45.

XX P-PSDB; AAW05178.

XX New isolated large tumour suppressor gene - used to develop prods.

XX for inhibiting cell proliferation or for enhancing proliferation

XX Claim 6; Page 118-123; 215pp; English.

XX This sequence encodes a human large tumour suppressor h-lats protein, and has been isolated from a fetal human brain phage lambda-gt10 cDNA library using a 2.1-kb DNA probe from the mouse m-lats gene (AAT42119). The full-length cDNA is present in plasmid pBS(KS)-h-lats.

XX The gene encodes a putative protein-serine/threonine-kinase, and inhibits cell proliferation and plays a crucial role throughout development. Activators or inhibitors of lats function (e.g. an antisense oligonucleotide or dominant-negative lats fragment) may be used in therapy of cancer or other proliferative disorders,

XX degenerative disorders, trauma, growth deficiency, etc., and fragments of the gene may be used as diagnostic probes. A

XX lats-inhibitor sequence may be expressed in a transgenic plant or

XX farm animal to confer increased growth and inhibit senescence.

XX Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;

XX Query Match 24.1%; Score 760; DB 17; Length 3984;

XX Best Local Similarity 63.7%; Pred. No. 6e-187;

XX Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

Qy 985 CAACCTGAGCCCTCACTCCCGCCCAACACGCTGACCGCGTGCAGCGCCGACACATC 1044

Db 1653 CACTCTGCTAATTCACGCTTCTGCTACAAAGTCACTGCAATTCACAGCTCTCTATT 1712

Qy 1045 CTTCAACCTGTGAAGAGCGTGTGCTGCGGCGCCGAGCCCAACAGACAGCGTGGGCCC 1104

Db 1713 CAACAGCCTGTGAAAGTATGCGTGTATTAACACAGAGCTACAGCTGCTTTAGCACT 1772

Qy 1105 TCGCACCCCGCTGGTGGCTGGCCCAACAGCCTGCCACTGAGAGCCTGGAGACGAG 1164

Db 1773 ACACACCCCTTCTTGATACCAACAGCCCAATTCAAACTGTTCACCCAGTCTTTCTCTAG 1832

Qy 1165 GAGGCGAGCGAGCCCAACACCCCGCTGATGTGGACTATGCGGCTCCGAGCGAGGTGC 1224

Db 1833 GGAACCGCTTCAATGTGACTGTGATGCCACTGTGCTGAAGCTCCAAATATCAAGGA 1892

Qy 1225 CCACCGCTTCGATATCCAAAGCACTTGTCTGCTGCCAGTAAAGTCTGAGCAGTACAGCGTG 1284

Db 1893 CCACCAACCCCTACCCAAACATCTGCTGCACCAAAACCCATCTGTTCTCCATACAG 1952

Qy 1295 GACCTGGACACCTGTGCACCAAGTGTGCAGCAGAGTCTGCGAGGGGCGCTGATCTAGAC 1344

Db 1953 TCAATCAGTAAGCTAGCAAAAGAGATCAGCAAGCTTGCCTCCCAAGGAAGATGAGAGTGA 2012

Qy 1345 GGGAGTGACAGAGCCCAAAAGGTGCGNAGGAGGACAAAGCTGGCAGAGACAAAAGCAG 1404

Db 2013 AAGAGTTA-----TGAAATGTTAGTGGGATAAAGAAAGAACAG 2057

Qy 1405 ATTCAGACCTCCCGGTGCTGTCGCAAGAAATAGCAGAGATGAAGAGAGAGAGTCT 1464

Db	3138	CACATTCACACACAAGCTAAACTCAGTCTCTGAGCTTCTGATCTTATTTATAACTTTGC	3197
Qy	2545	TGCGCGGCTGACTCGCGCCTGGCAGGGATGGGCGAGATGACCTCAAGGCACACCGCTTC	2604
Db	3198	CGAGACCCGAGATCGCTTAGGCAAGATGGTGATGATAAATAAAGCTCATCATTT	3257
Qy	2605	TTCAACACCATCGACTTTTCCGCTGACATCCGAAAGAGGCTGCACCTAGTCCCCACG	2664
Db	3258	TTTAAACAATTGACTTCTCGAGTGCACCTGAGACAGCAGTCTGCTCATACATTCCTAAA	3317
Qy	2665	ATCAGCACCCCATGGACACCTCCAAATTTGACCCGGTGGATGAAGAAAGCCCTGGCAC	2724
Db	3318	ATCACACACCAACAGATACATCAAATTTTCATCTGTTGATCCTGATAAATATGGAGT	3377
Qy	2725	GAGGCCAGC---GGAGAGAGCCNAGGCCCTGGGACACGCTGCGCTCCCCACAGCAGCA	2781
Db	3378	GATGATACGAGGAAGAAAATGTAANTGACACTCTCAATGGATGGTATAAATAATGGAAG	3437
Qy	2782	CATCCAGACAGCCCTCTTATGAGTTCACCTTCGCGAGGTTCTTCGATGACAACGCTAT	2841
Db	3438	CATCTCTGACATGCATTCTATGAATTTTACCTTCGGAAGGTTTTTGTGACAAATGGCTAC	3497
Qy	2842	CCCTTCGGGTGCCGAAGCC	2861
Db	3498	CCATATAATTATCGGAAGCC	3517

RESULT 14
AAZ51505
ID AAZ51505 standard; DNA; 3984 BP.

DT 21-JUN-2000 (first entry)

Homo sapiens.

```

21.1.3023
/*tag= a
/product= "Lats protein"

```

02-MAR-2000

F 18-AUG-1999; 99WO-US19068

1669500 1520, 380S-009699/ X

Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S; Turenchall CS

P-PSDB; AAY70390.

those UI latex proteins, -complete that is refractory to treatment

therapy, and disorders associ

XX Claim 44; Fig 12; 134pp; English.

CC The present sequence is a DNA encoding human Lats (large tumour
CC suppressor) protein which is a cell overproliferation inhibitor and a
CC negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
CC The present sequence is useful for treating cancer that is refractory
CC to standard chemotherapy or radiation therapy such as hyperplasia,
CC metaplasia, or dysplasia, and disorders associated with aberrant
CC levels of cdc2 activity. Conditions treated by promoting cdc2 function
CC include degenerative disorders, growth deficiencies, hypoproliferative
CC disorders, physical trauma, lesions, and wounds. An animal model
CC preferably a mouse, in which a lats gene has been disrupted by homologous
CC recombination, e.g. a lats knock-out mouse, is used for screening
CC compounds that can be used to treat or prevent cancer, particularly
CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.

XX SQ Sequence 3984 BP; 1280 A; 847 C; 798 G; 1059 T; 0 other;

Query Match 24.1%; Score 760; DB 21; Length 3984;
Best Local Similarity 63.7%; Pred. No. 6e-187;
Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

QY 985 CAACCTGAGCCCTACTGCCCCCCCCCAACACGCGTCACCGCGTGACGCGCCGACACATC 1044
DB 1653 CACTCTGCTAATTCACGCTTCTGCTACACAGCTACGCAATTTACACGCTCTCTATT 1712

QY 1045 CTTACCCCTGTGAAGAGCGTGGTGTGTCGGGCGGCGGACGACGCGCGTGGGGGCC 1104
DB 1713 CAACAGCCTGTGAAGATGTCGTATTAAACACAGAGCTACAGCTGCTTTAGCACT 1772

QY 1105 TCGCACCCGCGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1164
DB 1773 ACACACCTTCTTGATACACAGCCCAATTCAACTGTTCAACCCAGCTCTTTTCTCTAG 1832

QY 1165 GAGGACGCGGACGCGCCACACCGCTGGATGTGGAATGCGGCTCGGAGCGCAGGTGC 1224
DB 1833 GGAACCGCTTCAATGTGACTGTGATGCGACCTGTTGCTGAAGCTCCAACTATCAAGA 1892

QY 1225 CCACCGCTCGGTATCAAGAGCTTGTGTCGGCCAGTAAGTCTGACAGTACAGCGTG 1284
DB 1893 CCACACACCGCTACCCAAAACATCTGCTGCACCAAAACCCATCTGTTCTCCATACAG 1952

QY 1285 GACCTGGACAGCTGTGCACCAAGTGTGCAGCAGAGCTGCGAGGGGCGACTGATCTAGAC 1344
DB 1953 TCAATCACTAAGCTTAGCAAGAGGATCAGCCAACTTGGCCCAAGAGATGAGAGTGA 2012

QY 1345 GGGAGTGACAGAGCCACAAAGGTGCGAAGGAGACAAAGCTGGCAGACAAAAGCAG 1404
DB 2013 AAGAGTTA-----TGAAATGTTGATAGTGGGATAAAGAAAGAAACAG 2057

QY 1405 ATTACAGCTCCCGGTGCTGCGCAAGATACAGAGATGAGAGAGAGAGAGTCT 1464
DB 2058 ATTACAGCTACCTATTACTGTTAGAAAACAAGAAAGATGAAAGCGGAAGTCT 2117

QY 1465 CGCATCAAGAGTTACTCCCTTATGCTTCAATTTCTATGAGCAACACGCTGAGAAAT 1524
DB 2118 CGTATCAAGTTATCTCTCAAGCATTTAAATTTCTTTATGAGCAGCATGTAGAAAT 2177

QY 1525 GTCATCAAAAGCTACAGCAGAGGTGACCGCGAGGCTACAGCTGGCAGGAAATGCC 1584
DB 2178 GTACTCAAACTCTACAGCAGCGTCTACATCGTAAAAACAATTAGAAGTGAATGATG 2237

QY 1585 AAGCTGGGCTGTGAGCGCGCAGCAGGAGGATGAGAGATCCTCTACAGAGAGAG 1644
DB 2238 CGGTTGATTTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAA 2297

QY 1645 TCTAACTACAACCGCTGAAGAGGCCCAAGATGGACAGTCCATGTTGTGAAATCAAG 1704
DB 2298 TCTAATTAACCTGCTTAAAGGGCTAAATGGACAAAGTCTATGTTGTGAAGATAAG 2357

QY 1705 ACTCTAGGATCGGTGCGCTTTGGGGAAGTGTGCTCGCTTGTAAAGCTGGACACTCACGCT 1764
DB 2358 ACACCTAGGAATAGGAGCATTTGGTCAAGTCTGTAGCAAGAAAGTAGATACTAAGGCT 2417

QY 1765 CTGTAGCGCATGAAGACTCTCAGGAAGAGGATGCTTGAACCGGAATCAAGTGGCCCAT 1824
DB 2418 TTGTATGCAACAAAACCTCTTCAAGAAAGATGTTCTTTCGAAATCAAGTGCCTCAT 2477

QY 1825 GTCAAGGCTGAGAGGAGCATCTGCTGCTGAAGCAGACAAATGAGTGGTGGTCAACTCTAC 1884
DB 2478 GTTAAGGCTGAGAGAGATATCTGCTGAAGCTGACAAATGAATGGGTAGTGTCTATAT 2537

QY 1885 TACTCTCTCCAGGCAAGGACAGCCTGTACTTTGTGTGAGCTACATACAGCGCGGAT 1944
DB 2538 TATTCAATCCAAAGATAAGGACAATTTATCTTTGTGAATGAGCTACATCTCTGGGGTGT 2597

QY 1945 ATGATGAGCGCTGCTCATCAGGATGAGGCTCTCCCTGAGCAGCCTGCGCGCTTCTACAT 2004
DB 2598 ATGATGAGCGCTTATTAATTTAGAAATGGGCATCTTTCAGAAAGTCTGGCAGCAT 2657

QY 2005 GCAGAGTTGACCTGCGCATTTGAAAGTGTCCACAAAGATGGGCTTTATCCACCGGACATC 2064
DB 2658 GCAGAACTTACCTGTGAGTGAAGAGTGTTCATAAATGGGTTTATTCATAGAGATAT 2717

QY 2065 AAGCCTGACACATACTCATCGACCTGGATGCTCATATTAAGCTGACAGATTTTGGCCTC 2124
DB 2718 AAACCTGATATATTTTGTATGATGCTGATGCTCATATTAATTTGACTGACTTTGCGCTC 2777

QY 2125 TGCACTGATTCAGTGGACTCACAATTTCCAAAGTACTACGAAAGGAAACACATGAGA 2184
DB 2778 TGCACTGCTTCAGATGAGCAGACACGATTTCTAGTACTATCAGAGTGGTGCACCATCCAGG 2837

QY 2185 CAGGACAGCATGAGCGCGTGACCTCTGGGACGATGTTCCAACTGCTGCTGTGAGAC 2244
DB 2838 CAAGTAGATGAGATTTCAAGTAATGAATGGGGGATCCCTCAAGCTGCTGATGGAGAC 2897

QY 2245 AGTTTAAAGACCTCGGACAGAGCGCGGAGAGAGAGAGAGAGTGGCTGGCAGCATTTCT 2304
DB 2898 AGACTGAAGCCATTAGCGGGAGAGCTGCAGCCAGCAGCAGCATGCTAGCAGCATTTCT 2957

QY 2305 CTTGTGGGACACCAAAATTCATCTCGGAGGTGCTTCTCCGCAAGGGGTACAGGAG 2364
DB 2958 TTGGTTGGGACTCCCAATTTATTCACCTGAAGTGTGCTACGAACAGGATACACAG 3017

QY 2365 CTCTGCTGCTGGTGGAGCGTGGTGTGATCTCTTTGAGATGCTGTTGGGAGCGGCT 2424
DB 3018 TTGCTGATTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3077

QY 2425 TTCTTTGGGCCCCACCCCGCAGAGAGAGAGAGAGTGAAGGTGATCACTGGGAGAGCAGCTG 2484
DB 3078 TTCTTTGGCACAACACCATTTAGAAAACAATGAAGTTATCAACTGGCAACATCTCTT 3137

QY 2485 CATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCGCCGAGACCTCATCAGNAGCTGTC 2544
DB 3138 CAACTTCCACCAACAGCTAAAGCTAGTCTGAAAGCTTCTGATCTTATTTAAACTTTGC 3197

QY 2545 TCGCGGCTGACTGCGCGCTGGCGAGGATGGGCGAGATGACTCAAGCGACACCCGCTTC 2604
DB 3198 CGAGGACCCGAGATCGCTTAGGCAAGATGCTGCTGATGAATTAAGAGCTCATCTAT 3257

QY 2605 TTCAACACCATCGACTTTTCCCGTGACATCCGAAAGAGAGGCTGACCCCTAGCTCCCCACC 2664
DB 3258 TTAAACAAATTTGACTTCTCCAGTCACTGAGCAGCAGTCTGCTTATCATATTTCTAAA 3317

QY 2665 ATCAGCCACCCCATCGACACCTCCAAATTTTGACCCGGTGGATGAGAAAGCCCTTGGCAC 2724
DB 3318 ATCACACACCCCAACAGATACATAAATTTGATCTGCTGATCTTATTTAAATTTATGGAGT 3377

QY 2725 GAGGCGAGC---GGAGAGAGCGCCAAAGCTGGGACAGCTGGGCTCCCGCAGCAGCAAG 2781
DB 3378 GATGATACGAGGAGAGAAATGTAATGACACTCTCAATGGATGGTATTAATAATGGAAG 3437

QY 2782 CATCCAGAGCAGCGCTTCTATGAGTTTCACTTCCGCGAGGTTCTTCGATGACAAACGGCTAT 2841

This is the nucleotide sequence of a cDNA clone coding for a human orthologue, i.e. hWART1 (see AY405267), of Drosophila non-receptor serine/threonine kinase WART. hWART1 cDNA was isolated from a human bone marrow cDNA library using PCR fragments of hWART1 as probes. 3 Polymorphisms were identified in the hWART1 gene: (1) at nucleotide 978, resulting in an Ala/Gly change; (2) at nucleotide 1840, silent; and (3) at nucleotides 3252-3253, comprising a deletion of 2 adenosines, resulting in a C-terminal truncation of hWART1 in the putative kinase domain. The latter frameshift mutation was observed in 2 independent clones from human bone marrow cDNA. Truncation of hWART1 could play a role in disease progression. hWART1 shows strong expression in cell lines from non-small cell lung cancer, ovarian tumours, central nervous system tumors, renal tumours and breast tumours, and may provide a target for oncology drug development. Nucleic acids encoding full-length hWART1 and hWART1 polypeptides lacking one or more of amino acid segments 12-45, 55-151, 236-377, 404-530, 555-559, 601-702, 691-998 and 1011-1086, or lacking one or more of the N-terminal domain, catalytic domain, or C-terminal domain are claimed, as well as hWART2 sequences (AA87397), hWART1 and hWART2 polypeptides, antibodies a method for identifying modulators of hWART function, and use of such modulator compounds to treat an abnormal condition and use of such modulator compounds to treat an abnormal condition involving hWART signal transduction, especially cancer. Probes for hWART1 and hWART2 nucleic acids are also claimed.

```

CC - detection of HWART HUCLEIC ACIDS
XX
SQ      Sequence 7382 BP; 2291 A; 1458 C; 1417 G; 2216 T; 0 other;
      Query Match      24.1%; Score 760; DB 20; Length 7382;
      Best Local Similarity 63.7%; Pred. No. 8.2e-187;
      Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps
      2
QY      985 CAACCTGAGCCCTCACTGCGCGCCCAACAGCGTCAACGCGTGACGGCGGCACACATC 1044
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1817 CACTCTGCTAATTCCTCAGCCCTTCTGTACACAGTCACTGCAATTAACACCAGCTCTATT 1876
QY      1045 CTTCAACCCCTGTGAAGACGCTGCGTGTGCGGCGGAGCCCCAGACGCGTGGGGCCC 1104
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1877 CAACAGCCTGTGAAAAGTATGTCGTGTATTAACAAACAGGAGCTACAGCTGCTTTASCACCT 1936
QY      1105 TCGCACCCCGCTGGGTGCTGGCGCCACACAGCACTTGCACCTGAGAGCGTGGAGACGAG 1164
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      1937 ACACACCCCTCTTGGATACACAGCCCAATTCAAACTGTTCAACCCAGTCTTTTCTCTGAG 1996
QY      1165 GAGGCGACGCGAGGCGCCACACCCGCTGGATGTGGAGCTATGTGGCGGCTCCAGCGCAGGTGC 1224
Db      1997 GGAACCGCTTCAAATATGTGACTGTGATGCCACTTGTGCTGAAGCTCCAAACTATCAAGGA 2056
QY      1225 CCACCGCTTCGCTATCCAAAGCACTTCTGTCGTGCCCAAGTAACTGTGACGAGTACAGCGTG 1284
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2057 CCACACCAACCTACCCAAAACATCTGTGCACCAAAACCCATCTGTCTCTCCATCAGAG 2116
QY      1285 GACCTGGACAGCCTGTGCACCAGTGTGCACGAGAGTCTGGCAGGGGGCACTGTATCTAGAC 1344
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2117 TCAATCAGTAAAGCTTAGCAAGAGGATCAGCAAGCTTGCCCAAGGAAGATGAGAGTGAA 2176
QY      1345 GGGAGTGCACAAGAGCCACAAGAGTTCGAAGGGAGACAACAGCTGCCACAGACAACAGCAG 1404
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2177 AAGAGTTA-----TGAATAATGTTGATGTGGGGATTAAGAAAGAAACAG 2221
QY      1405 ATTCAGACCTCCCGGTGCTGTCCGCAAGATAGCAGAGATGAAGAGAGACAGAGAGTCT 1464
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2222 ATTACAACTTCACCTATTACTGTTAGGAAAAACAAGAAAGATGAAGAGCGAAGGGAATCT 2281
QY      1465 CGCATCAAGAGTTTACTCCCGTTATGCCTTCAAATTTTCATGTGAGCAACACGCTGGAGAA 1524
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2282 CGTATTCAAAGTATTCTCTCAAGCATTTAAATTTCTTTATGGAGCAACATGTAGAAAA 2341
QY      1525 GTCATCAAAACCTTACCAGCAGAAAGGTTCAGCCGAGGCTTACGCTGGACGAGGAATGCC 1584
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db      2342 GTACTCAAAATCTCATCAGCAGCGTCTACATCGTFAAAAAACAATTAGAAGATGAAGATG 2401
QY      1585 AAAGCTGGGCTCTGTAGGCCGAGGAGGACGATCAGCAACATCTCTTACCAGAGGAG 1644
      ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

```

DB		3438	CATCCTGAACATGCATTTCATAAATTTACCTTCCGAGGTGTTTTTGATGACAATGGCTAC	3497
QY		2842	CCCTTCGGTGCCCGAAGCC	2861
DB		3498	CCATAAATTAATCCGAAGCC	3517
RESULT 15				
ID		AAx87396	standard; cDNA; 7382 BP.	
XX		AAx87396;		
DD		08-OCT-1999	(first entry)	
DT				
XX			Human WART1 cDNA.	
DE				
KW		WART1; hWART1; WART orthologue; human; signal transduction;		
KW		protein kinase; cancer; tumour; diagnosis; therapy; ss.		
XX		Homo sapiens.		
OS				
XX		Key	Location/Qualifiers	
FF		5'UTR	1..399	
FT			/*tag= a	
FT		repeat_region	12..63	
FT			/*tag= b	
FT			/note= "contains 10 copies of GGC repeat, similar	
FT			to repeats that undergo expansion in human	
FT			diseases associated with neuronal	
FT			phenotypes"	
FT		CDS	395..3787	
FT			/*tag= c	
FT		3'UTR	3788..7382	
FT			/*tag= d	
FT		repeat_region	6058..6346	
FT			/*tag= e	
FT			/note= "Alu-J subfamily repeat motif"	
FT		polyA_signal	7339..7344	
FT			/*tag= f	
FT		variation	replace(978,G)	
FT			/note= "results in Ala/Gly change"	
FT		variation	replace(1840)	
FT			/*tag= h	
FT			/note= "silent polymorphism"	
FT		variation	replace(3252..3253)	
FT			/*tag= i	
FT			/note= "deletion of 2 adenines, causes	
FT			frameshift"	
XX		WO9937787-A2.		
PN				
XX		29-JUL-1999.		
XX			99WO-USO1145.	
XX		20-JAN-1999;		
PF				
XX		21-JAN-1998;	98US-0072023.	
PR				
XX		(SUGE-) SUGEN INC.		
PA				
XX		Flanagan P, Plowman GD;		
PI				
DR		WPI; 1999-458698/38.		
DR		P-PSDB; AAU06326.		
XX		New nucleic acid encoding human orthologs of Drosophila WART		
PT		proteins, used to identify specific modulators for treating cancer		
PT		or for diagnosis		
XY				

claim 1! page 120-121: 137pp; English.

Db 2402 CCGGTTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATCCTTTGCCAAAAGAA 2461

Qy 1645 TCTAACTACACCGGCTGAAGAGGCCAAGATGGACAAAGTCCATGTTGTGAAAATCAAG 1704

Db 2462 TCTAATTACATCCGCTTAAAGAGGCTAAATGGACAGTCTATGTTGTGAAGATAAG 2521

Qy 1705 ACTCTAGGACCGGTGCGCTTTGGGGAAGTGGCTGCTGTTGAAGCTGGACACTCACGGT 1764

Db 2522 ACACCTAGGAATAGGAGCATTTGGTGAAGTCTGCTAGCAAGAAAGTAGATACTAAGGCT 2581

Qy 1765 CTGTAGCGCATGAAGACATCTCAGGAGAGAGGATGCTGCAACCGGAATCAAGTGCCCAT 1824

Db 2582 TTGTATGCAACAAAACCTCTCGAAGAAAGATGTTCTTCGAAATCAAGTCGCTCAT 2641

Qy 1825 GTCAAGGCTGAGAGGACATCTCGCTGAAGCAGACAAATGAGTGGTGGTCAAACTCTAC 1884

Db 2642 GTTAAGGCTGAGAGATATCTCGCTGAAGCTGACATGAATGGTAGTTCGCTATAT 2701

Qy 1885 TACTCCTTCCAGGACAAAGACAGCCTGTACTTGTGATGGACTACATACCAAGCGGGAT 1944

Db 2702 TATTCATTCCAAGATAAGACAAATTTATCTTGTAAATGGACTACATTCCTGGGGTGAT 2761

Qy 1945 ATGATGAGCCTCTCATCAGGATGAGGCTCTCCCTGAGCACCTGGCCCGCTTCTACATT 2004

Db 2762 ATGATGAGCCTATTAATAGAAATGGCATCTTTCCAGAAAGTCTGGCACGATTTCTACATA 2821

Qy 2005 GCAGAGTTGACCTGGCCTGAAAGTGTCCACAGATGGGCTTTATCCACGGGACATC 2064

Db 2822 GCAGAACTTACCTGTGAGTGAAGATGTTTATAAATGGGTTTATTCATAGAGATATT 2881

Qy 2065 AAGCCTGACAACTACTCATCGACCTGGATGCTCATATTAAGCTGACAGATTTGGCCCTC 2124

Db 2882 AAACCTGATATATTTTGATGATCGTGTGCTATATTAATGACTGACTTTGSCCTC 2941

Qy 2125 TGCACGTGATTCAGGTGACCTACAAATTCCTAGTACTACAGAGAGGAAACACATGAGA 2184

Db 2942 TGCACGTGCTTCAGATGGACACACAGATTTCTAAGTACTACAGAGTGGTGAACATCCACGG 3001

Qy 2185 CAGGACAGCATGGAGCCGGTGACCTCTGGAGCATGTTTCCAACTGCTGCTGGAGAC 2244

Db 3002 CAAGATAGATGGATTCAGTAATGAATGGGGGATCCCTCAAGCTGCTGATGCTGGAGAC 3061

Qy 2245 AGTTTAAAGCCTGGAGAGCGGCGAGAGACAGCAGAGGCTGCTGGCAGCATTTCT 2304

Db 3062 AGACTGAAGCCATTAGAGCGGAGAGCTGCACGCCACACAGCGATGCTAGCACATCT 3121

Qy 2305 CTGTCGGGACACCAAAATTAATCCTCCGAGGCTGCTCCCGCAAGAGGTACAGCAG 2364

Db 3122 TTGGTTGGGACTCCCAATTAATATGCACTGAAGTGTGCTACGAAACAGGATACACAG 3181

Qy 2365 CTCGTGACTGCTGAGGCTGCTGATCTCTTTGAGATGCTGTTGGGAGCGGCT 2424

Db 3182 TTGCTGATTTGGTGGATGTTGGTGTATCTTTTGAATGTTGGTGGGACCACTCT 3241

Qy 2425 TTTTGGCCCCACCCACACAGAGCGAGCTGAAGTGAATCAACTGGGAGAGACACGCTG 2484

Db 3242 TTTTGGCACAAACACCAATTAAGAAACAAATGAAGTTATCAACTGGCAAACTCTCT 3301

Qy 2485 CATATCCTAGCAGGTGAGGCTAGCGCTGAGGCCCCAGACCTCATCAGAAAGCTGTGC 2544

Db 3302 CACATTCACCAACAGCTAAACTCAGTCTGAAGCTTCTGATCTTATTTAAACTTTGC 3361

Qy 2545 TGGCGGCTGACTCGGCTGGGAGGATGGGAGATGACCTCAAGGCACACCCGTTTC 2604

Db 3362 CGAGACCCGAGATCGGTTAGGAAAGATGGTCTGATGAATAAAAGCTCATCTATT 3421

Qy 2605 TTCAACACCATCGACTTTTCCGCTGACATCCGAAAGAGGCTGCACCTAGCTCCACC 2664

Db 3422 TTTAAACAAATGACTTCTCCAGTACCTGAGACAGAGCTGCTGTATACATTCCTAAA 3481

Qy 2665 ATCAGCCACCCATGGACACCTTCAATTTTACCGGCTGGATGAAGAAAGCCCTGGCAC 2724

Db 3482 ATCACACACCAACAGATACATCAAAATTTTGATCCTGTTGATCCTGATAAATATGGAGT 3541

Qy 2725 GAGGCCAGC---GGAGAGAGCGCCCAAGGCTGGGACACGCTGGCCTCCCCAGCAGCAAG 2781

Db 3542 GATGATAACGAGGAGAAAATGTAATGACACTCTCAATGGATGATATAAAAATGGAAG 3601

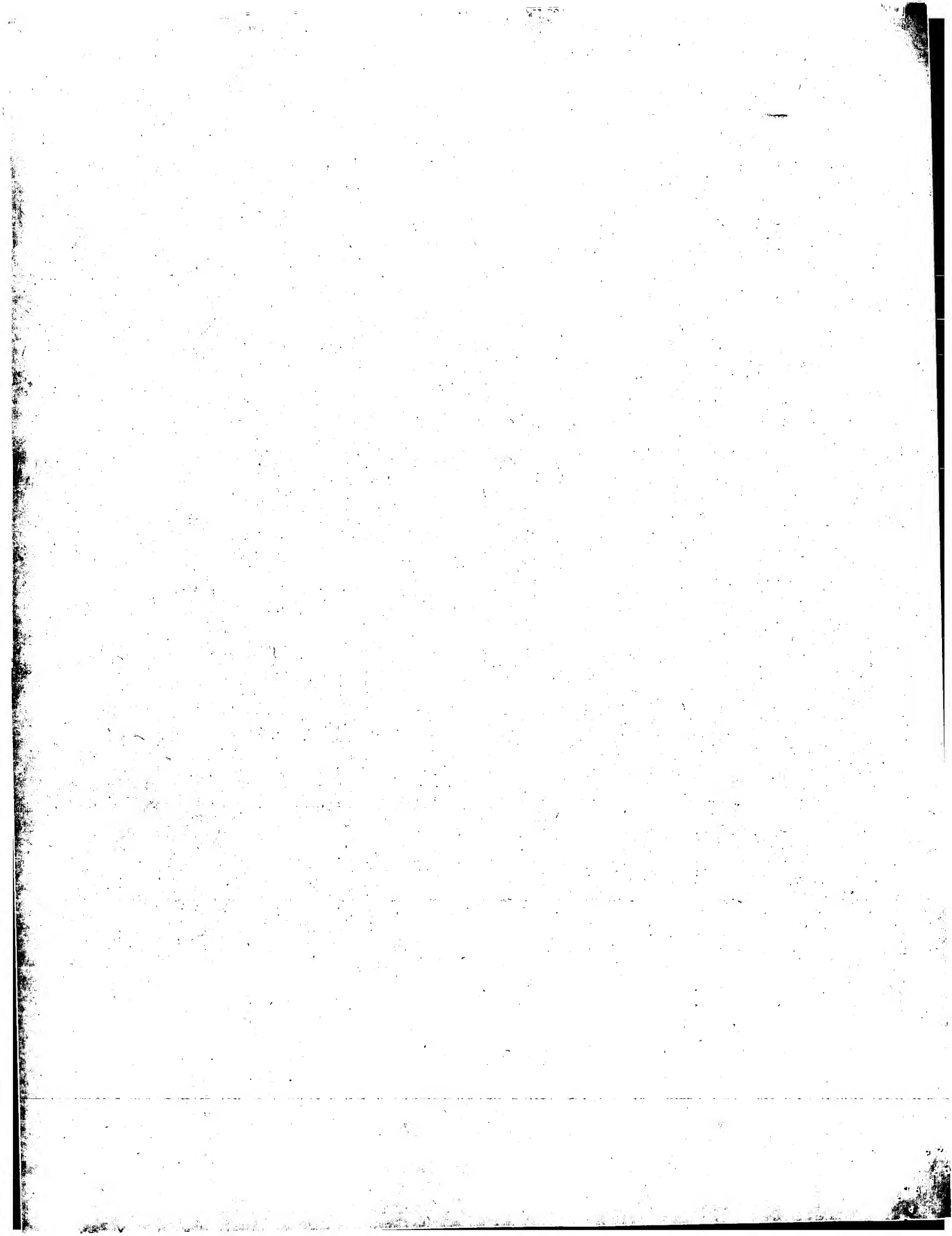
Qy 2782 CATCCAGAGCAGCGCTTCTATGAGTTTCCAGGCTTCCGAGGTTCTTCGATGACAAACGGCTAT 2841

Db 3602 CATCCTGAACATGCAATCTATGAATTTACCTTCCGAAGGTTTTTTGATGACAATGGCTAC 3661

Qy 2842 CCCTTCCGCTGCCCGAAGCC 2861

Db 3662 CCATATAATATCCGAAGCC 3681

Search completed: January 16, 2003, 10:34:06
Job time : 464.423 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:02 ; Search time 64.1915 seconds
(without alignments)
15073.116 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagaccacccgaagtt.....aagagcattatttggggg 3155

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA:*

- 1: /cgn2_6/ptodata/1/ina/5A.COMB.seq:*
- 2: /cgn2_6/ptodata/1/ina/5B.COMB.seq:*
- 3: /cgn2_6/ptodata/1/ina/6A.COMB.seq:*
- 4: /cgn2_6/ptodata/1/ina/6B.COMB.seq:*
- 5: /cgn2_6/ptodata/1/ina/PCRTUS.COMB.seq:*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	3155	100.0	3155	4 US-09-442-100-7	Sequence 7, Appli
2	1293	41.0	1961	4 US-09-509-902A-15	Sequence 15, Appli
3	1148	4	1498	4 US-09-509-902A-6	Sequence 6, Appli
4	812.6	25.8	3213	4 US-09-442-100-5	Sequence 5, Appli
5	760	24.1	3984	4 US-09-442-100-3	Sequence 3, Appli
6	564.6	17.9	5720	4 US-09-442-100-1	Sequence 1, Appli
7	333	10.6	638	4 US-09-328-111-26	Sequence 26, Appli
8	258.8	8.2	2160	4 US-09-588-256-1	Sequence 1, Appli
9	233.8	7.4	1935	2 US-08-878-989-11	Sequence 11, Appli
10	233.8	7.4	1935	2 US-03-272-796-11	Sequence 11, Appli
11	225.2	7.1	3018	2 US-08-860-150-6	Sequence 6, Appli
12	225.2	7.1	3018	2 US-03-338-132-6	Sequence 6, Appli
13	222.6	7.1	2101	2 US-08-860-150-1	Sequence 1, Appli
14	222.6	7.1	2101	3 US-09-338-132-1	Sequence 1, Appli
15	214.4	6.8	678	4 US-09-328-111-66	Sequence 1, Appli
16	161.8	5.1	2549	4 US-09-467-082-3	Sequence 66, Appli
17	149.8	4.7	2726	2 US-08-422-699A-12	Sequence 3, Appli
18	149.8	4.7	2726	2 US-08-422-706B-12	Sequence 12, Appli
19	144.4	4.6	3182	1 US-08-484-044-11	Sequence 12, Appli
20	143.6	4.6	1599	3 US-09-256-465-1	Sequence 1, Appli
21	143.6	4.6	1599	4 US-09-167-322-3	Sequence 3, Appli
22	143.4	4.5	2511	2 US-08-422-699A-8	Sequence 8, Appli
23	143.4	4.5	2511	2 US-08-422-706B-8	Sequence 8, Appli
24	126.4	4.0	2610	2 US-09-212-771-1	Sequence 1, Appli
25	126.4	4.0	2610	3 US-09-091-058-1	Sequence 1, Appli
26	123.8	3.9	4739	3 US-08-685-871-1	Sequence 1, Appli
27	121.8	3.9	5053	2 US-08-685-576-2	Sequence 2, Appli

28	120.2	3.8	4363	2 US-08-685-576-5	Sequence 5, Appli
29	118.8	3.8	2104	4 US-09-313-930-1	Sequence 1, Appli
30	106.8	3.4	1732	4 US-09-430-564-1	Sequence 1, Appli
31	102.2	3.2	2206	1 US-08-221-817-10	Sequence 10, Appli
32	102.2	3.2	2206	1 US-08-454-439-10	Sequence 10, Appli
33	102.2	3.2	2206	5 PCT-US94-10487-10	Sequence 10, Appli
34	102.2	3.2	2848	4 US-08-464-954A-2	Sequence 2, Appli
35	100.6	3.2	2204	1 US-08-221-817-12	Sequence 12, Appli
36	100.6	3.2	2204	1 US-08-454-439-12	Sequence 12, Appli
37	100.6	3.2	2204	5 PCT-US94-10487-12	Sequence 12, Appli
38	99	3.1	2557	4 US-08-464-954A-1	Sequence 1, Appli
39	97.4	3.1	1637	2 US-08-966-316-10	Sequence 10, Appli
40	95.8	3.0	1890	3 US-09-289-466-2	Sequence 2, Appli
41	95.8	3.0	1891	3 US-09-289-466-1	Sequence 1, Appli
42	95.8	3.0	1929	2 US-09-016-000-10	Sequence 10, Appli
43	94.4	3.0	3323	2 US-08-422-699A-10	Sequence 10, Appli
44	94.4	3.0	3323	2 US-08-422-706B-10	Sequence 10, Appli
45	93	2.9	2599	6 5266464-1	Patent No. 5266464

ALIGNMENTS

RESULT 1
US-09-442-100-7
Sequence 7, Application US/09442100
Patent No. 6359193
GENERAL INFORMATION:
APPLICANT: Xu, Tian
APPLICANT: Tao, Wufan
APPLICANT: Wang, Weiyl
APPLICANT: Zhang, Sheng
APPLICANT: Yu, Wan
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
TITLE OF INVENTION: GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mistrock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3155 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: CDNA
FEATURE:
NAME/KEY: CDS
LOCATION: 1..2943

US-09-442-100-7

Query Match		100.0%	Score 3155;	DB 4;	Length 3155;		
Best Local Similarity		100.0%	Pred. No. 0;				
Matches 3155;	Conservative 0;	Mismatches 0;	Indels 0;	Gaps 0;			
QY	1	ATGAGAGCCACCCGAGTTTGGACCTTATCAAAAAGCTCTCAGGGAATTCGATATTC	60				
Db	1	ATGAGAGCCACCCGAGTTTGGACCTTATCAAAAAGCTCTCAGGGAATTCGATATTC	60				
QY	61	CTCCTGCCCTTTTGGCAACAGTTCAGGCACTTCGGCAGCTGCAGAGTGAACCGGCAGATG	120				
Db	61	CTCCTGCCCTTTTGGCAACAGTTCAGGCACTTCGGCAGCTGCAGAGTGAACCGGCAGATG	120				
QY	121	CTTCAGGAGTTGGTGAATCGCGCATGTGACAGAGATGGCTGGCAGAGCGCTCACGCAG	180				
Db	121	CTTCAGGAGTTGGTGAATCGCGCATGTGACAGAGATGGCTGGCAGAGCGCTCACGCAG	180				
QY	181	ACGGCAGTAGAGTATCGAAGCTGCCCTGGAGTACATCAGTAAGATGGCTACCTGGAC	240				
Db	181	ACGGCAGTAGAGTATCGAAGCTGCCCTGGAGTACATCAGTAAGATGGCTACCTGGAC	240				
QY	241	CCCAGGAATGAGCAGATTTGCGAGTCTCAAGCAGACCTCCCGAGGAAGGGCTGGCG	300				
Db	241	CCCAGGAATGAGCAGATTTGCGAGTCTCAAGCAGACCTCCCGAGGAAGGGCTGGCG	300				
QY	301	TCCACCCCGGTGACTCGCGGGCCAGTTTCGAGGGCAGAGGGGAAGCACTCCCATCTAC	360				
Db	301	TCCACCCCGGTGACTCGCGGGCCAGTTTCGAGGGCAGAGGGGAAGCACTCCCATCTAC	360				
QY	361	CACCAGCTGGGTGGTGCAAACTACGAGGGCCCGCGCACTGGAGAGATGCCCGGCAA	420				
Db	361	CACCAGCTGGGTGGTGCAAACTACGAGGGCCCGCGCACTGGAGAGATGCCCGGCAA	420				
QY	421	TATTTAGACTTTCTTCCCTGGAGCGGAGCGGCAACCGAGCTGCCAGGCTCACAG	480				
Db	421	TATTTAGACTTTCTTCCCTGGAGCGGAGCGGCAACCGAGCTGCCAGGCTCACAG	480				
QY	481	CATCTCTCCCAAGGGTACAGCAGCAGTAGAGCCAAGTGGCACTTTCGGGGCACAC	540				
Db	481	CATCTCTCCCAAGGGTACAGCAGCAGTAGAGCCAAGTGGCACTTTCGGGGCACAC	540				
QY	541	TATGGTCGTGGTCTACTATCGAGCAGTCTGGGTATGGGTGACGCGCAGTTCTCTCC	600				
Db	541	TATGGTCGTGGTCTACTATCGAGCAGTCTGGGTATGGGTGACGCGCAGTTCTCTCC	600				
QY	601	TTCCAGAACAGAGCCACAGATGCTTATCCAGCATGGCCAGGCCAGGGTGGCCCT	660				
Db	601	TTCCAGAACAGAGCCACAGATGCTTATCCAGCATGGCCAGGCCAGGGTGGCCCT	660				
QY	661	CCGCCAGGCTTACCTTTCTGCGCATGCTGGGTGTACACTGCTCGCACCAAGCCG	720				
Db	661	CCGCCAGGCTTACCTTTCTGCGCATGCTGGGTGTACACTGCTCGCACCAAGCCG	720				
QY	721	CGGCTTACCCACCTTGGGGCCACCACTTACATGTGTGGGCAACCGGGTCCACGTTT	780				
Db	721	CGGCTTACCCACCTTGGGGCCACCACTTACATGTGTGGGCAACCGGGTCCACGTTT	780				
QY	781	ACTGGGCAAGCTTGCACAGGCTGTGCTGGCACCGTCCAGGAACAGCCTCAATGTGAC	840				
Db	781	ACTGGGCAAGCTTGCACAGGCTGTGCTGGCACCGTCCAGGAACAGCCTCAATGTGAC	840				
QY	841	TTGTACGAGCTGGGCTCCAGGCTGCCCTGGTCTGCAGCTCCACTGGCAGCGCGACTCG	900				
Db	841	TTGTACGAGCTGGGCTCCAGGCTGCCCTGGTCTGCAGCTCCACTGGCAGCGCGACTCG	900				
QY	901	CTGCAGAGAGGGTCTAGAAGCCTTCGCGCGCGCATGTGGCTTTTGGGCTGGCCCCAGC	960				
Db	901	CTGCAGAGAGGGTCTAGAAGCCTTCGCGCGCGCATGTGGCTTTTGGGCTGGCCCCAGC	960				
QY	961	AGGACCAACTCTTCAACAAACCAACCTGAGCCCTCACTGCGCGCCCCAACAGGTC	1020				
Db	961	AGGACCAACTCTTCAACAAACCAACCTGAGCCCTCACTGCGCGCCCCAACAGGTC	1020				

QY	1021	ACGCCGTGACGGCGGCACACATCTTTCACCTGTGAAGAGCGTGGCTGCTCGGGCC	1080				
Db	1021	ACGCCGTGACGGCGGCACACATCTTTCACCTGTGAAGAGCGTGGCTGCTCGGGCC	1080				
QY	1081	GAGCCCCAGACAGCGCTGGGGCCCTCGCACCCCGCTGGGTGGCTGGCCACAGCACCT	1140				
Db	1081	GAGCCCCAGACAGCGCTGGGGCCCTCGCACCCCGCTGGGTGGCTGGCCACAGCACCT	1140				
QY	1141	GCCACTGAGAGCTGGAGACGAAGAGGGCAGCGAGGCCACACCCCGCTGGATGTGAC	1200				
Db	1141	GCCACTGAGAGCTGGAGACGAAGAGGGCAGCGAGGCCACACCCCGCTGGATGTGAC	1200				
QY	1201	TATGGCGCTCCGAGCGAGTGGCCACCGCTCCGCTATCAAGACACTTGTGCTGCC	1260				
Db	1201	TATGGCGCTCCGAGCGAGTGGCCACCGCTCCGCTATCAAGACACTTGTGCTGCC	1260				
QY	1261	AGTAAGTCTGAGCAGTACAGCTGGACCTGCACAGCTGTGCACAGTGTGCAGCAGAT	1320				
Db	1261	AGTAAGTCTGAGCAGTACAGCTGGACCTGCACAGCTGTGCACAGTGTGCAGCAGAT	1320				
QY	1321	CTGGAGGGGCACTGATCTAGACGGGAGTGACAAGAGCCACAAAGTGCAGAGGAGAC	1380				
Db	1321	CTGGAGGGGCACTGATCTAGACGGGAGTGACAAGAGCCACAAAGTGCAGAGGAGAC	1380				
QY	1381	AAAGCTGGCAGACACAAAAGCAGATTACAGACCTCCCGCTGCTGCCCAAGATAGC	1440				
Db	1381	AAAGCTGGCAGACACAAAAGCAGATTACAGACCTCCCGCTGCTGCCCAAGATAGC	1440				
QY	1441	AGAGTGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTTCAAAATTC	1500				
Db	1441	AGAGTGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTTCAAAATTC	1500				
QY	1501	TTTCATGAGCAACACCTGGAGATGTCATCAAAACCTACAGAGAGGTACGCGGAGG	1560				
Db	1501	TTTCATGAGCAACACCTGGAGATGTCATCAAAACCTACAGAGAGGTACGCGGAGG	1560				
QY	1561	CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTGAGCCGAGCAGAGCAGATG	1620				
Db	1561	CTACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTGAGCCGAGCAGAGCAGATG	1620				
QY	1621	AGGAAGCTCTTACAGAGGAGTCTAATAACACCGCTGAAGAGGSCCAAGATGAC	1680				
Db	1621	AGGAAGCTCTTACAGAGGAGTCTAATAACACCGCTGAAGAGGSCCAAGATGAC	1680				
QY	1681	AAGTCCATGTTGTGAAAATCAAGACTCTAGCATCGGCTTTGGGAAAGTGTGCCTC	1740				
Db	1681	AAGTCCATGTTGTGAAAATCAAGACTCTAGCATCGGCTTTGGGAAAGTGTGCCTC	1740				
QY	1741	GCTTGTAGCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGGAAGAGATGTC	1800				
Db	1741	GCTTGTAGCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGGAAGAGATGTC	1800				
QY	1801	CTGAACCGGAATCAAGTGGCCCATGTCAAGCTGAGAGGAGATCCTGGCTGAAGCAGAC	1860				
Db	1801	CTGAACCGGAATCAAGTGGCCCATGTCAAGCTGAGAGGAGATCCTGGCTGAAGCAGAC	1860				
QY	1861	AATGAGTGGGTGGTCAAACTCTACTCTCTTCCAGGACAAAGGACAGCTGTACTTTGG	1920				
Db	1861	AATGAGTGGGTGGTCAAACTCTACTCTCTTCCAGGACAAAGGACAGCTGTACTTTGG	1920				
QY	1921	ATGGACTACATACAGCGGGGATATGATGACCTTGTGATCAGGATGAGGCTCTTCCT	1980				
Db	1921	ATGGACTACATACAGCGGGGATATGATGACCTTGTGATCAGGATGAGGCTCTTCCT	1980				
QY	1981	GAGCACTTGGCCGCTTCTACATTCAGAGTGTACCTTGGCCATTAAGATGTCCACAG	2040				
Db	1981	GAGCACTTGGCCGCTTCTACATTCAGAGTGTACCTTGGCCATTAAGATGTCCACAG	2040				
QY	2041	ATGGCTTTTATCCACCGGGACATCAAGCTGACAACATCTATCTGACCTGGATGGTCAT	2100				
Db	2041	ATGGCTTTTATCCACCGGGACATCAAGCTGACAACATCTATCTGACCTGGATGGTCAT	2100				

QY 2101 ATTAAGTACAGAGATTTGGGCTCTGCACTGGAGTTCAGTGGACTCACAATTCCAAGTAC 2160
Db 2101 ATTAAGTACAGAGATTTGGGCTCTGCACTGGAGTTCAGTGGACTCACAATTCCAAGTAC 2160
QY 2161 TACCAGAAAGGACACATGACAGAGACAGCATGAGCCCGGTGACCTCTGGGACGAT 2220
Db 2161 TACCAGAAAGGACACATGACAGAGACAGCATGAGCCCGGTGACCTCTGGGACGAT 2220
QY 2221 GTTTCACACTGCTGCTGTGGAGACAGTTTAAAGACCTCTGAGCAGAGGCGGAGAGCAG 2280
Db 2221 GTTTCACACTGCTGCTGTGGAGACAGTTTAAAGACCTCTGAGCAGAGGCGGAGAGCAG 2280
QY 2281 CACAGAGTGGCTGGACATCTCTGTGCGGACACCAATATACATCCTCCGGAGGTG 2340
Db 2281 CACAGAGTGGCTGGACATCTCTGTGCGGACACCAATATACATCCTCCGGAGGTG 2340
QY 2341 CTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAGGCTGGGTGTGATCTCTTT 2400
Db 2341 CTTCTCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAGGCTGGGTGTGATCTCTTT 2400
QY 2401 GAGATGCTGTTGGGACGCGCTTCTTGGGCGCCACCCACACAGAGCAGCTGAAG 2460
Db 2401 GAGATGCTGTTGGGACGCGCTTCTTGGGCGCCACCCACACAGAGCAGCTGAAG 2460
QY 2461 GTGATCAACTGGGAGAGCAGCTGTGATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCC 2520
Db 2461 GTGATCAACTGGGAGAGCAGCTGTGATATCCCTACGAGGTGAGGCTCAGCGCTGAGGCC 2520
QY 2521 CGAGACCTCATCAGNAAGTGTGCTGCGGCGGTGACTGCGCGCTGGGAGGAGTGGGCA 2580
Db 2521 CGAGACCTCATCAGNAAGTGTGCTGCGGCGGTGACTGCGCGCTGGGAGGAGTGGGCA 2580
QY 2581 GATGACCTCAAGGACACACCGCTTCTTCAACACCATGCTGCTTCCCGTGCACATCCGAAG 2640
Db 2581 GATGACCTCAAGGACACACCGCTTCTTCAACACCATGCTGCTTCCCGTGCACATCCGAAG 2640
QY 2641 CAGGCTGACCTTACGCTCCCGACCATCAGCCACCCCATGAGACCTTCCAAATTTGACCGG 2700
Db 2641 CAGGCTGACCTTACGCTCCCGACCATCAGCCACCCCATGAGACCTTCCAAATTTGACCGG 2700
QY 2701 GTGGATGAAGAAAGCCCTGGACAGGCGCCAGGAGAGAGCGCCAGGCTGGGACAG 2760
Db 2701 GTGGATGAAGAAAGCCCTGGACAGGCGCCAGGAGAGAGCGCCAGGCTGGGACAG 2760
QY 2761 CTGGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTACCTTCCGACG 2820
Db 2761 CTGGGCTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTACCTTCCGACG 2820
QY 2821 TTCTTGGATGACACGCGCTATCCCTTCCGCTGCCGAGCCCTCAGAGCCCGCAGAGAT 2880
Db 2821 TTCTTGGATGACACGCGCTATCCCTTCCGCTGCCGAGCCCTCAGAGCCCGCAGAGAT 2880
QY 2881 GCAGACCCAGGGATGCGGACTTGGAGGTGGGCGGAGGCTGCCAGCGGTGTACGTG 2940
Db 2881 GCAGACCCAGGGATGCGGACTTGGAGGTGGGCGGAGGCTGCCAGCGGTGTACGTG 2940
QY 2941 TAAGCTCAGTTTAAACACAACTCGAGGAAACCCAAATTCAGATTTCTTTTCAAGAGACAA 3000
Db 2941 TAAGCTCAGTTTAAACACAACTCGAGGAAACCCAAATTCAGATTTCTTTTCAAGAGACAA 3000
QY 3001 ACTCAAGCTTAGGAATCCCTTCAATTTTGTGTTAAATGGGCAACAGAGAGTCAAC 3060
Db 3001 ACTCAAGCTTAGGAATCCCTTCAATTTTGTGTTAAATGGGCAACAGAGAGTCAAC 3060
QY 3061 ATGATTTCAATTAGCCCTCTGAGGACCTTCACTGCAATTAACACAGTATTTTAAAAA 3120
Db 3061 ATGATTTCAATTAGCCCTCTGAGGACCTTCACTGCAATTAACACAGTATTTTAAAAA 3120
QY 3121 TTAGTACAGTATGAAAGAGCAGCTATTTTGGGG 3155
Db 3121 TTAGTACAGTATGAAAGAGCAGCTATTTTGGGG 3155

RESULT 2
US-09-509-902A-15
; Sequence 15, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 15
; LENGTH: 1961
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-15

Query Match 41.0%; Score 1293; DB 4; Length 1961;
Best Local Similarity 84.4%; Pred. No. 0;
Matches 1497; Conservative 0; Mismatches 255; Indels 21; Gaps 3;

QY 1185 CCCGCTGGATGTGGACTATGCGGCTCCGAGCGCAGGTGCCCCACCGCCTCCCGTATCCCAA 1244
Db 2 CCCGCTGGAGTGGAGTACGGAGGCCACGCGGAGGTCCCGCTCCGCCCTACCCGAA 61
QY 1245 GCACCTGTGCTGCCCGAGTAACTCTGAGCAGTACAGCGTGGACCTGGACAGCTGTGCAC 1304
Db 62 GCACCTGTGCTGCCCGAGTAACTCTGAGCAGTACAGCGTGGACCTGTGCAC 115
QY 1305 CAGTGTGACGAGAGTCTGCGAGGGG---CACTGATCTAGAGGGAGTGACAAGACA 1361
Db 116 AGGCAATGGAGCAGAGCTCCGTCGCGGCCCAACAGCGCGGCGGCGACAAGACGG 175
QY 1362 CAAAGTGGGAAGGAGCAAAAGCTGGCAGAGACAAAAGAGATTTCAGACCTCCCGGT 1421
Db 176 CAAAGTGGGAAGGAGCAAAAGCTGGCAGAGACAAAAGAGATTTCAGACCTCCCGGT 235
QY 1422 GCCTGTGCGCAAGAATAGCAGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTTATC 1481
Db 236 TCCGCTCCGCAAAAACAGCAGAGAGAGAGAGAGAGTCAAGAGTACTC 295
QY 1482 CCCTTATGCTTCAATTTCTTCATGAGCAACAGCTGGAGATGTCATCAAAACCTACCA 1541
Db 296 GCATACGCTTTAAGTTCTTCATGAGCAGCAGCTGGAGATGTCATCAAAACCTACCA 355
QY 1542 GCAGAGGTTCAGCGGAGGCTACAGTGGAGCAGGAAATGGCAAAGCTGGGCTCTGTGA 1601
Db 356 GCAGAGGTTCAGCGGAGGCTACAGTGGAGCAGGAAATGGCAAAGCTGGGCTCTGTGA 415
QY 1602 GCGCAGCAGGAGCAGATGAGGAGATCCTCTACAGAGAGAGTCTAACTACAACCGGT 1661
Db 416 AGCTGAGCAGGAGCAGATGCGGAAGATCCTCTACAGAAAGAGTCTAATTAACAAGGTT 475
QY 1662 GAAGAGGCCAAGATGGACAAGTCCATGTTGTGAAATCAAGACTCTAGGCTCGGTGC 1721
Db 476 AAAGAGGCCAAGATGGACAAGTCTATGTTGTCAAGATCAAAACCTGGGATCGGTGC 535
QY 1722 CTTTGGGGAAGTGTGCTCGCTTGTAAAGTGGACACTCACGCTGTAGCGCATGAAGAC 1781
Db 536 CTTTGGGGAAGTGTGCTCGCTTGTAAAGTGGACACTCACGCTGTAGCGCATGAAGAC 595
QY 1782 TCTCAGGAAGAGATGCTGTAACCGGAATCAAGTGGCCCCCATGTCAAGGCTGAGAGGA 1841
Db 596 CTTAAGAAAAAGGATGCTCTGAACCGGAATCAGGTGGCCCCACGTCAGAGCGGAGGGA 655
QY 1842 CATCTGCTGAAGCAGACATGAGTGGTGTAAACTTACTACTCTACTCTCTCTCCAGACA 1901
Db 656 CATCTGCGCGGAGGACAGCAATGAGTGGTGTAAACTTACTACTCTCTCTCTCCAGACA 1901

```

QY 1902 GCACAGCCCTGACTTTGTGATGAGTACATACACGCGGGGATATGATGAGCCTGCTGAT 1961
Db 716 AGACAGCCCTGACTTTGTGATGAGTACATACCTGCTGGGACATGATGAGCCTGCTGAT 775
QY 1962 CAGGATGAGGCTTCCCTGAGCAGCTGGCCGCTTCTACATTCGACAGTTGACCCCTGGC 2021
Db 776 CCGGATGAGGCTTCCCTGAGCAGCTGGCCGCTTCTACATTCGACAGTTGACCTTGGC 835
QY 2022 CATTGAAAGTGTCCACAGATGGCTTTATCCACCGGACATCAAGCCTCAGACATAC 2081
Db 836 CATTGAGAGTGTCCACAGATGGCTTTATCCACCGGACATCAAGCCTCAGACATAC 895
QY 2082 CATGACCTGATGCTATATTAAGCTGACAGATTTTGGCCTGTGACCTGATTCAGGTG 2141
Db 896 GATAGATCTGATGCTGATTAAGCTGACAGATTTTGGCCTGTGACCTGATTCAGGTG 955
QY 2142 GACTCACAATTTCCAAAGTACTACCAAGAGGACACACATGAGACAGACAGATGAGCC 2201
Db 956 GACTCACAATTTCCAAATATTACCAAGAGGACACATGAGACAGACAGATGAGCC 1015
QY 2202 CGGTGACCTCTGGGACGATTTTCCAACTGCTGTGGAGACAGGTTAAAGACCTTGA 2261
Db 1016 CAGCGACCTCTGGGATGATGTCTAACTGCTGGTGTGGGACAGGCTCAAGACCTTGA 1075
QY 2262 GCAGAGGCGCGACAGCAGCAGAGGTCCTGGCAGATTTCTTGTGCGGACACAAA 2321
Db 1076 GCAGAGGCGCGACAGCAGCAGAGGTCCTGGCAGATTTCTTGTGCGGACACAAA 1135
QY 2322 TTACATGCTCTGGGACGATTTTCCAACTGCTGTGGAGACAGGTTAAAGACCTTGA 2381
Db 1136 CTACATGCTCTGGGACGATTTTCCAACTGCTGTGGAGACAGGTTAAAGACCTTGA 1195
QY 2382 CGTCGCTGATTTCTTTCAGATGCTGTGGGACGCTGTGACCTGCTGAGTGTGAG 2441
Db 1196 TGTGAGTGTATTTCTTTCAGATGCTGTGGGACGCTGTGACCTGCTGAGTGTGAG 1255
QY 2442 CACAGAGCGCAGCTGAAGGTGATCAACTGGGAGACAGCTGCTGATATTCCTACGAGGT 2501
Db 1256 CACAGAGCGCAGCTGAAGGTGATCAACTGGGAGACAGCTGCTGATATTCCTACGAGGT 1315
QY 2502 GAGGCTGAGCTGAGGCGGAGCTGATCACTGAGGAGACAGCTGCTGATATTCCTACGAGGT 2561
Db 1316 GAGGCTGAGCTGAGGCGGAGCTGATCACTGAGGAGACAGCTGCTGATATTCCTACGAGGT 1375
QY 2562 CTTGGGAGGATGGGCGGAGCTGATCACTGAGGAGACAGCTGCTGATATTCCTACGAGGT 2621
Db 1376 CTTGGGAGGATGGGCGGAGCTGATCACTGAGGAGACAGCTGCTGATATTCCTACGAGGT 1435
QY 2622 TTCCCGTGCATCCGAAAGAGGCTGACCTGAGGAGACAGCTGCTGATATTCCTACGAGGT 2681
Db 1436 TTCCCGTGCATCCGAAAGAGGCTGACCTGAGGAGACAGCTGCTGATATTCCTACGAGGT 1495
QY 2682 CACCTCCAAATTTGACCGGCTGATGAGGAGGCTGACCTGAGGAGACAGCTGCTGATATTCCTACGAGGT 2741
Db 1496 CACCTCCAAATTTGACCGGCTGATGAGGAGGCTGACCTGAGGAGACAGCTGCTGATATTCCTACGAGGT 1555
QY 2742 CGCCAGGCTTGGGACAGCTGGCTTCCCGGAGCAGCAGCAGCTGAGGAGGCTTGA 2801
Db 1556 CACCAAGGCTTGGGACAGCTGGCTTCCCGGAGCAGCAGCAGCTGAGGAGGCTTGA 1615
QY 2802 TGAGTTCACCTTCCCGGAGGCTTCCCGGAGCAGCAGCAGCTGAGGAGGCTTGA 2861
Db 1616 CGAATTCACCTTCCCGGAGGCTTCCCGGAGCAGCAGCAGCTGAGGAGGCTTGA 1675
QY 2862 CTCAG-----AGCCCGCAGAGTGTGACAGCAGGAGGCTGAGGAGGCTTGA 2909
Db 1676 TTGAGGAGCAGAGGCTTTCAGAGCTGAGGAGGCTGAGGAGGCTTGA 1735
QY 2910 TGCGGCGGAGGCTTTCAGAGCTGAGGAGGCTGAGGAGGCTTGA 2942
Db 1736 TCAGACTGAAGGCTTTCAGAGCTGAGGAGGCTGAGGAGGCTTGA 1768

```

```

RESULT 3
US-09-509-902A-6
; Sequence 6, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:
; APPLICANT: Virca, Duke
; APPLICANT: Bird, Timothy A.
; APPLICANT: Anderson, Dirk M.
; APPLICANT: Marken, John S.
; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
; FILE REFERENCE: 2877-US
; CURRENT APPLICATION NUMBER: US/09/509,902A
; CURRENT FILING DATE: 1999-08-03
; NUMBER OF SEQ ID NOS: 16
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 1498
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-509-902A-6

```

```

Query Match 36.4%; Score 1148.4; DB 4; Length 1498;
Best Local Similarity 86.3%; Pred. No. 8.6e-293;
Matches 1295; Conservative 0; Mismatches 196; Indels 9; Gaps 2;
QY 1185 CCCGCTGGATGTGGAGTATCGCGGCTCCGAGCGCAGGTGCGCCACCGCTCCGCTATCCAAA 1244
Db 4 CCCGCTGGAGTGTGGAGTATCGCGGCTCCGAGCGCAGGTGCGCCACCGCTCCGCTATCCAAA 63
QY 1245 GCATCTGCTGCTGCCAGTAACTGTGAGCAGTACAGCGTGGACCTGGACAGCTGTGAC 1304
Db 64 GCACCTGCTGCTGCCAGTAACTGTGAGCAGTACAGCGTGGACCTGGACAGCTGTGAC 117
QY 1305 CAGTGTGACAGAGTGTGCGAGGGG---CACTGATCTAGACGGGAGTGCAGAGGCA 1361
Db 118 AGCATGTGAGCAGAGCTGCGTGGCGGCCCAAGAGCGCGGCGGCGAGCAGAGCG 177
QY 1362 CAAAGGTGCGAGGAGACAAAGCTGGCAGAGACAAAGAGTTCAGACTCCCGCT 1421
Db 178 CAAAGGTGCGAGGAGACAAAGCTGGCAGAGACAAAGAGTTCAGACTCCCGCT 237
QY 1422 GCTGTGCGCAGAGTGTGAGAGTGAAGAGAGAGAGTCTCGCATCAGAGTACTC 1481
Db 238 TCCGCTGCGCAGAGTGTGAGAGTGAAGAGAGAGTCTCGCATCAGAGTACTC 297
QY 1482 CCTTATGCTTCAATTTCTTCTATGAGGACACAGTGTGAGAGTGTCAATCAACAG 1541
Db 298 GCATACGCTTTAAGTTCTTCTATGAGGACACAGTGTGAGAGTGTCAATCAACAG 357
QY 1542 GCAGAGTGTGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1601
Db 358 GCAGAGTGTGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 417
QY 1602 GCGCAGGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 1661
Db 418 AGCTGAGGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 477
QY 1662 GAAAGGCGCAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 1721
Db 478 AAAGAGGCGCAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 537
QY 1722 CTTTGGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1781
Db 538 CTTTGGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 597
QY 1782 TCTCAGGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGT 1841
Db 598 CTTTGGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 657
QY 1842 CATCTGCTGAGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 1901
Db 658 CATCTGCTGAGAGGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAGAGTGTGAG 717

```

QY 1902 GGACAGCCTGTACTTTGTGATGACTACATACAGCGGGGATATGATGAGCCTGTGTAT 1961
 Db 718 ACACAGCCTGTACTTTGTGATGACTACATCCTGTGGGACATGATGAGCCTGTGTAT 777
 QY 1962 CAGGATGAGGCTTCCCTGACACCTGGCCGCTTCTACATGAGAGTTGACCTGGC 2021
 Db 778 CCGGATGAGGCTTCCCTGACACCTGGCCGCTTCTACATGAGAGTTGACCTGGC 837
 QY 2022 CATTGAAGTGTCCCAAGATGGCTTTATCCACGGGACATCAAGCCTGACCAACTAT 2081
 Db 838 CATTGAGAGTGTCCCAAGATGGCTTTATCCACGGGACATCAAGCCTGATCAACTTT 897
 QY 2082 CATGACCTGGATGTGATTAATGATGACAGATTTTGGCCCTCTGACCTGGATTCAGTG 2141
 Db 898 GATGATGCTGGATGTGATTAATGATGACAGATTTTGGCCCTCTGACCTGGATTCAGTG 957
 QY 2142 GACTCACAATTCAGTACTACCAAGAGGGAACACATGACAGGACGATGAGCC 2201
 Db 958 GACTCACAATTCAGTACTACCAAGAGGGAACACATGACAGGACGATGAGCC 1017
 QY 2202 CGGTGACCTCTGGGACGATTTTCCAACTGTGCTGTGGAGACAGTTAAAGACCTTGA 2261
 Db 1018 CAGCAGCTCTGGGATGTGCTTAAGTGTGGGACAGCTGAAGACCTTGA 1077
 QY 2262 GCAGAGGGCGGAGAGCAGCAGAGGTGCTTGGCAGATTTCTTGTGGGACACCAA 2321
 Db 1078 GCAGAGGGCGGAGAGCAGCAGAGGTGCTTGGCAGATTTCTTGTGGGACACCAA 1137
 QY 2322 TTACATGCTCGGAGGTGCTTCTCGCAAGGGTACAGGAGCTGTGACTGTGGAG 2381
 Db 1138 TTACATGCTCGGAGGTGCTTCTCGCAAGGGTACAGGAGCTGTGACTGTGGAG 1197
 QY 2382 COTCGGTGTGATCTTCTTGTGATGCTGTGGGAGCGCTTCTTGGCCCGCCACCC 2441
 Db 1198 GTTGGAGTGTCTTCTGAGATGCTGTGGGAGCGCTTCTTGGCCCGCCACCC 1257
 QY 2442 CACAGAGCGCAGCTGAAGTGTATCACTGGGAGAGCAGCTGTGATATCCCTTACGAGT 2501
 Db 1258 CACAGAAACCCAGCTGAAGTGTATCACTGGGAGAGCAGCTGTGATATCCCTTACGAGT 1317
 QY 2502 GAGGCTCAGCTGAGGCGGAGAGCTTATCAGAGCTGTGCTGGGCGCTGATGCGG 2561
 Db 1318 GAAGTGAAGCTGAGGCGGAGAGCTTATCAGAGCTGTGCTGGGCGCTGATGCGG 1377
 QY 2562 COTGGGAGGAGTGGGAGAGTGTATCACTGAAGCAGCAGCTTCTTCAACACCTGACTT 2621
 Db 1378 COTGGGAGGAGTGGGAGAGTGTATCACTGAAGCAGCAGCTTCTTCAACACCTGACTT 1437
 QY 2622 TTCCTGTGATCTCGAAAGCAGGCTGACCTTACCTTACCTTACCTTACCTTACCTT 2681
 Db 1438 CTCAGTGTACATCCGAAAGCAGGCTGACCTTACCTTACCTTACCTTACCTTACCTT 1497

RESULT 4
 US-09-442-100-5

; Sequence 5, Application US/09442100
 ; Patent No. 6359193

GENERAL INFORMATION:

; APPLICANT: Xu, Tian
 ; APPLICANT: Tao, Wufan
 ; APPLICANT: Wang, Weiyl.
 ; APPLICANT: Zhang, Sheng
 ; APPLICANT: Yu, Wan
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711

COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent in Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/442.100
 FILING DATE:

CLASSIFICATION:

PRIOR APPLICATION DATA: 08/411.111

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Mirostock, S. Leslie

REGISTRATION NUMBER: 18,872

REFERENCE/DOCKET NUMBER: 6523-003

TELECOMMUNICATION INFORMATION:

TELEPHONE: (212) 790-9090

TELEFAX: (212) 869-9741/8864

TELEX: 66141 PENNIE

INFORMATION FOR SEQ ID NO: 5:

SEQUENCE CHARACTERISTICS:

LENGTH: 3213 base pairs

TYPE: nucleic acid

STRANDEDNESS: double

TOPOLOGY: unknown

MOLECULE TYPE: CDNA

FEATURE:

NAME/KEY: CDS

LOCATION: 1..2889

US-09-442-100-5

Query Match 25.8%; Score 812.6; DB 4; Length 3213;

Best Local Similarity 65.1%; Pred. No. 3.4e-204;

Matches 1237; Conservative 0; Mismatches 644; Indels 18; Gaps 2;

QY 985 CAACCTGAGCCCTACTGCCGCCCAACACGCTGACCGCTGACGCGCCGACACATC 1044
 Db 919 CACTCTGTCTTCTCAGCCCTTCTGCCACTCAGTCCATCACACCCCTCTATT 978
 QY 1045 CTTACCTCTGTAAGAGCGTGGTGTGCTGCGGCCGCGCCAGCCGACAGCCGCTGGGCGC 1104
 Db 979 CAACAGCCGCTGAAAGCATGCGCTCTCTGAAACACAGAGCTCGAGACTGCTTTAGCCCA 1038
 QY 1105 TCGCACCCGCTGCTGCTGCGCCACAGCAGCTGCTGAGAGCTGAGAGCTGAGAGAG 1164
 Db 1039 ACCATCTCTTGTGATGCCACAGCCAGTTCAGACTGTTGAGCTTACCCCTTTTCTGAG 1098
 QY 1165 GAGGAGCAGCAGGCGCCACACCCGCTGATGTGGACTATGGCGCTCCGAGCGCAGGTGC 1224
 Db 1099 GGTACAGCTTCAAGTGTGCTGTATCCACCTGTGCTGAAGCTCCAAGCTATCAAGT 1158
 QY 1225 CCACCGCTCCGTATCCAAAGCAGTTCCTGCTGCCAGTAACTGAGCAGTACAGCGTG 1284
 Db 1159 CCACCGCTCCGTATCCAAAGCAGTTCCTGCTGCCAGTAACTGAGCAGTACAGCGTG 1218
 QY 1285 GACCTGGAGCAGCTGTGCTGACAGCTGTGAGAGCTGCGAGGCGGCGACTGACTAGAC 1344
 Db 1219 TCAGTAACTGAGCCCTGCAAGATGCAAGCTTACCCCAAGAGATGATAGTACAG 1278
 QY 1345 GGGAGTGAAGAGCCCAAGAGTTCGGAAGGAGGAGCAAGCTGGCAGAGCAAAAAGCAG 1404
 Db 1279 AAGAGTGGGACAGT-----GGTGACTCTGGGATATAAGAAAGAAACAG 1323
 QY 1405 ATTCAGACTTCCCGGTGCTGTCGCAAGATATACAGAGATGAAGAGAGAGAGTCT 1464
 Db 1324 ATTACACTTCACTTATCAGTGTTCGGAAGGAGGAGGAGGAGGAGGAGGAGGAGTCT 1383
 QY 1465 CGCATCAAGAGTACTTCCCTTATGCTTCAATTTCTTATGAGGAGCAACAGTGGAGAT 1524
 Db 1384 CGGATCAGGTTACTCTCCACAGGCTTTAAGTTCTTTTCTGAGGAGCAGTAGAGAC 1443
 QY 1525 GTCATCAAAACCTTACCAGCAAGAGGTCAGCGCGGAGGCTACAGCTGGAGCAAGAAATGGCC 1584

Db 1444 GTCCCTGAAGTCTCATCAGCAGCGCTCTGCATCGGAAGAGCAGCTAGAAAATGAAATGATG 1503
QY 1585 AAAAGTGGGCTCTGTGAGGCCGAGCAGGAGAGATGAGGAAGATCCTCTACCAGGAAGGAG 1644
Db 1504 CGGGTGGATATTCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAGAAAGAG 1563
QY 1645 TCTAATCTACACCGGCTGAAGAGGCCAAGATGGAAGTCCATGTTTGTGAAATCAAG 1704
Db 1564 TCTAATCTATTTCGCTTTAAAAGGCTAAAATGGACAAGTCTATGTTTGTAAAGATAAG 1623
QY 1705 ACTTAGCATCGGTCCTTTGGGAACTGTGCTCGCTTTGTAAGCTGGACACTCAGCT 1764
Db 1624 ACATTAGAATAGAGCGTTGGTGAAGTCTGTCTAGCAAGAAAGTCTGACTAAAGCT 1683
QY 1765 CTGTACGCCATGAAGACTCTCAGGAAGAGGATGTCTTGAACCGGAATCAAGTGGCCCAT 1824
Db 1684 TTGTATGCAACAAGACTCTTCGAAGAAAGACGTTCTGTCTCCGAATATCAGGTGGCTCAT 1743
QY 1825 GTCAAGGCTGAGAGGACATCCTGGCTGAAGCAGACAAATAGTGGTGGTCAAACTCTAC 1884
Db 1744 GTGAAGCGGAGAGGGATATCTAGCAGAACCGACAAATAGTGGTGGTCCGCTGTAC 1803
QY 1885 TACTCTTCCAGGACAAGCAGCTCTACTTGTGTGATGACTACATACAGGCGGGAT 1944
Db 1804 TACTCTTCCAGGACAAGCAGCTCTACTTGTGTGATGACTACATACAGGCGGGAT 1863
QY 1945 ATGATGACCTGTGATCAGGATGAGGCTCTTCCCTGAGCAGCTGGCCGCTTCTACATT 2004
Db 1864 ATGATGACCTTATTAATAGATGGCATCTTCTCGAATCTGGCAGCATCTACATA 1923
QY 2005 GCAGATGACCTGGCCATTGAAGTGTCCAAAGATGGCTTTATCCACCGGACATC 2064
Db 1924 GCAGAACTTACCTGTGAGTGAAGTGTTCATATAAATGGGTTTTATTATAGATATT 1983
QY 2065 AAGCTGCAACATACATCATGACCTGGATGGTTCATATTAGCTGACAGATTTGGCTC 2124
Db 1984 AACCTGATACATTTGATGACCTGTGAGGCTGATGGCCATTAATAATGACTGACTTTGGCTG 2043
QY 2125 TGCCTGATGAGTGGTGAATTCACAAATTCACAAAGTACTACCAAGAGGAACCATGAGA 2184
Db 2044 TGCCTGATGAGTGGTGAATTCACAAATTCACAAAGTACTACCAAGAGTGGGATCACCCAGG 2103
QY 2185 CAGCAGCAGTGGAGCGGCTGACCTCGGACAGATTTCCAACTGCTGCTGTGGAGAC 2244
Db 2104 CAAGTACATGATTTTCAGTAAGAAATGGGAGATCCTTCCAAATTTGCTGGTGGGAG 2163
QY 2245 AGTTAAAGACCTTGGACGAGAGCGGCGCAGAACGACAGCAGAGGTGCTTGGCAGCATTTCT 2304
Db 2164 AGCTGAAGCCACTGGAGCGGAGCTGCTGCCAGCAGCAGGATGTCTAGCCCATTTCT 2223
QY 2305 CTGTGCGGACCAAAATTCATGCTCGGAGGTGCTTCCGCAAGAGGTACAGCAG 2364
Db 2224 CTGTTGGGACTCCCAATTAATGCACTGAAAGTGTCTACTGCAAGAGATATACACAG 2283
QY 2365 CTCTGTGACTGGGAGCGCTGCTGATCTCTTTGAGATGCTGTTGGGAGCGGCT 2424
Db 2284 CTGTGACTGGGAGTGTGCTGTTTCTTTGTGAATGTTGTTGGGAGCAACCTCT 2343
QY 2425 TTCTTGGCCCCCCCCCAGAGAGCGGCTGAAGTGTATCAACTGGGAGAGCAGCTG 2484
Db 2344 TTCTTGGCACAACCCCAATTAGAACAACAAATGAAGTTATCATCTGGCAAACTTCTCTA 2403
QY 2485 CATATCCCTACGAGTGGCTGACGCTGAGGCCGAGAGCTCATCAGAGCTGTGC 2544
Db 2404 CATATCCCTCTCAAGCTAAGCTGAGTGTGCTGAGGCTCTGAGCTCATCAACTGTGT 2463
QY 2545 TGGCGGCTGACTGCGGCTTGGGAGGATGGGAGATGACCTCAAGCAGCAGCGGCTC 2604
Db 2464 CGAGGACCAAGAGACCGCTCGCAGAAAGCGGTGCTGATGAGATAAAGGCTCATCATTT 2523
QY 2605 TTCAACACCATCGACTTTTCCGCTGACATCCGAAAGCAGGCTGCACCTTACCTGCCACC 2664

Query Match 24.1%; Score 760; DB 4; Length 3984;
Best Local Similarity 63.7%; Pred. No. 2.6e-190;

Db 2524 TTTAAGACCATCGATTTCTTAGTGTATCTGAGACAGCAGTCTGCTTCATACATCCCTAAA 2583
QY 2665 ATCAGCCACCCATGGACACCTCCAAATTTTACCCGGTGGATGAGAAAGCCCTGGCAG 2724
Db 2584 ATCAGCATCCCAACAGATACATCAATTTGACCCCTGTGATCTCTGATTAATTTGGAGC 2643
QY 2725 GAGCCACAGCGAGAG---AGGCCCAAGCCCTGGGACACGCTGGCCTCCCCCAGCAGCAAG 2781
Db 2644 GATGCGCAGGAGGAGAAATATCAGTGCACACTCTGAGCGGATGGTATATAAATGGGAAG 2703
QY 2782 CATCCAGGACGCGCTTCTATGATTCACCTTCGCGAGGTTCTTCGATGACACGCTAT 2841
Db 2704 CACCCGAGCAGCGCTTCTATGATTCACCTTTCCGAGGTTTTTTGATGACAAATGGCTAC 2763
QY 2842 CCCTTCGCTGCCGAGAGCCCTCAGAGCCCGCAGAGAGT 2880
Db 2764 CCATATAATTATCCAAAGCCCTATTGAGTATGAATACATT 2802

RESULT 5

US-09-442-100-3
; Sequence 3, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiye
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; APPLICANT: Pennie & Edmonds
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; NUMBER OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESS: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER-READABLE FORM:
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent In Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442,100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411,111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mirock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 3984 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 231..3623
; US-09-442-100-3

Matches 1197; Conservative 0; Mismatches 665; Indels 18; Gaps 2;

QY 985 CAACCTGAGCCTCACTGCCCCCCCCCAACAGCGTCAACGCGTCAACGCGCCGACACATC 1044
 Db 1653 CACTCTGCTAAATCTCAGCGCTTCTGCTACACAGTCACTGCAATACACCGCTCCTATT 1712
 QY 1045 CTTCACTCTGTAAGAGCGTGTGCTGCTGCGCGCCGAGCCGACAGCGCTGGGGCC 1104
 Db 1713 CAACAGCCTGTGAAAGTATGCTGTATTAACACAGAGCTACAGACTGCTTTAGCACCT 1772
 QY 1105 TCGCACCGCGCTGGTGTGCTGCGCCACAGACACCTGCGCACTGAGAGCTGAGACGAG 1164
 Db 1773 ACACACCTCTTGGATACCAAGCAATTAACACTGTTCAACCCAGTCTTTTCTGTAG 1832
 QY 1165 GAGGAGCAGCGGAGCGCCACACCGCTGGATGTGGACTATGCGCGCTCCGAGCGCAGGTGC 1224
 Db 1833 GGAACCGCTTCAATGTGACTGTGATGCCACCTGTTGCTGAAGCTCCAACTATCAAGGA 1892
 QY 1225 CCACCGCTCTGATATCCAAAGCACTTGTGCTGCGCCAGTAAGTCTGAGCAGTACAGCGTG 1284
 Db 1893 CCACCAACCGCTTACCCCAACATCTGCTGCACCAAAACCCATCTGTTCTCCATACGAG 1952
 QY 1285 GACCTGGACAGCTGTGCACCACTGTGCAGCAGAGTCTGCAGGGGCGCACTGATCTAGAC 1344
 Db 1953 TCAATCAGTACCTAGCAAGAGGATCAGCCAGCTTGCCCAAGAGAGATGAGAGTAA 2012
 QY 1345 GGGAGTGAAGAGCCCAAGAGTGGAGGAGACAAAGCTGGCAGAGACAAAAGCAG 1404
 Db 2013 AAGAGTTA-----TGAATGTTGATAGTGGGATTAAGAAAGAAACAG 2057
 QY 1405 ATTCAGACTCCCGGTGCTGCGCAAGATAGCAGATGAAGAGAGAGAGTCT 1464
 Db 2058 ATTACAACCTTCACTTACTGTTAGGAAAACAAAGAGATGAAGAGCGAAGGAACT 2117
 QY 1465 CCATCAAGAGTTACTCCCTTATCCCTTCAATTTCTCATGTCAGCAACACGTGAGAT 1524
 Db 2118 CATTCAAGTTATCTCTCAAGCATTTAAATTTCTTTTGGAGCAACATGAGAAAT 2177
 QY 1525 GTCACTCAAAACCTACAGCAAGGTGACCGGAGGTACAGCTGGAGCAGGAAATGGCC 1584
 Db 2178 GTACTCAATCTCATCAGCAGGTCTACATCGTAAACAACTTAGAATGAATGATG 2237
 QY 1595 AAAGCTGGCTGTGTAGCGCAGCAGGAGCAGATGAGGAAGATCCTCTACAGAAAGG 1644
 Db 2238 CGGTTGATATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAA 2297
 QY 1645 TCTAACTACAACCGCTGAAGAGGCGCAAGATGGACAGTCCATCTTCTGGAATCAAG 1704
 Db 2298 TCTAATTCATCCGCTTTAAAGGGCTAAATGGACAGTCTATGTTTGTGAAGATAAG 2357
 QY 1705 ACTTAGCATCGGTGCTTTGGGGAAGTGTGCTCGCTTGTAACTGGACACTCACGCT 1764
 Db 2358 ACATPAGGAATAGGAGCATTTGGTGAAGTCTGTAGCAAGAAAGTAGATAAAGCT 2417
 QY 1765 CTGTAGCCATGAAGACTCTCAGGAAGAGATGTCCTGAACCGGAATCAAGTGGCCCAT 1824
 Db 2418 TTGTATGAACAAAACCTCTCGAAAGAAAGATGTTCTTCTCGAAATCAAGTGCCTCAT 2477
 QY 1825 GTCAAGGTGAGGAGGACATCTGCTGAAGCAGCAATGAGTGGGTGCTCAACTCTAC 1884
 Db 2478 GTTAGGCTGAGAGAGATATCTGCTGAAGCTGACAAATGAATGGTGTGCTGTATAT 2537
 QY 1885 TACTCTTCCAGGCAAGCAGCAGCTGTACTTGTGATGGACTACATACAGCGGGGAT 1944
 Db 2538 TATTCAATCCAAAGATAAGCAATTTATCTTGTATGAGTACTATCTTCTGGGGGTGAT 2597
 QY 1945 ATGATGAGCTGTGATCAGGATGAGGTCTTCCCTGAGCAGCTGCGCCGCTTCTACATT 2004
 Db 2598 ATGATGAGCTTATTAATTAAGATGGCATCTTCCAGAAAGTCTGGCAGCATTTACATA 2657
 QY 2005 GCAGATTTGACCTGGCCATGAAGTGTCCACAGATGGCTTTATCCACGGGACATC 2064
 Db 2658 GCAGAACTTACCTGTGCAGTTGAAAGTGTTCATAAAATGGGTTTTTATTATAGATATT 2717

QY 2065 AAGCCTGACAACATACTCATGACCTGGATGCTATATTAAAGTGAACATTTTGGCCTC 2124
 Db 2718 AAACCTGATAATATTTGATTGATCCTGATGTCATATTAAATGACTGACTTTGGCCTC 2777
 QY 2125 TCACTGGATTCAGGTGGACTCACAATTCAAAGTACTACAGAAAGGGAACACATAGAGA 2184
 Db 2778 TCACTGGCTTCAGATGGACACACAGATCTAAGTACTATCAGAGTGTGACCATCCACGG 2837
 QY 2185 CAGGACAGCATGGAGCCCGCTGACCTCTGGGAGGATGTTTCAACTGTGCTGTGGAGAC 2244
 Db 2838 CAAGATAGCATGGATTTCAGTAATGAATGGGGGATCCCTCAAGCTGTGCTGTGGAGAC 2897
 QY 2245 AGTTTAAACACCTCGAGCAGAGGCGCAAGACAGCAGCAGCAGAGTGCCTGGCACTTCT 2304
 Db 2898 AGACTGAAGCAATAGAGCGGAGAGCTGCACCCAGCAGCAGGATGTCTACACATCTT 2957
 QY 2305 CTTTGGGACACCAAAATTTACATCCTCGGAGGTCTTCTCCGAAAGGTTACAGCAG 2364
 Db 2958 TTGGTTGGGACTTCCCAATTTATTTGACCTGAAGTGTGCTACGAACAGGATACACAG 3017
 QY 2365 CTCTGCTGCTGCTGAGCGTCTGCTGATCTCTTTGAGATGCTGTTGGGAGCGGCT 2424
 Db 3018 TTGTGTGATTTGGTGGAGTGTGCTGTTAFTCTTTTGAATGTTGGTGGGCAACCTCT 3077
 QY 2425 TTCTTGGCCCCCACCACACAGAGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTG 2484
 Db 3078 TTTCTTGGCACAACACCATTTAGAAACACAAATGAAGTTATCACTGGCAACATCTCTT 3137
 QY 2485 CATATCCCTTACCGAGTGGAGCTCAGCGCTGAGCGCCGAGAGCTCATCACGAAGTGTG 2544
 Db 3138 CACATTTCCACCAAGCTAAAGCTCAGTCTGAGGCTTCTGATCTTATTATTAACCTTGC 3197
 QY 2545 TCGCGGCTGACTCGCGCTGGGAGGATGGGCGAGATGAGCTCAAGGCACACCGCTTC 2604
 Db 3198 CGAGGACCCGGAAGTCTGCTTAGGCAAGATGCTGTGATGAAATAAAAGCTCATCATTT 3257
 QY 2605 TTCAACACCATCGACTTTTCCGTCGACATCCGAAAGAGCTGACACCTTACCTCCACCC 2664
 Db 3258 TTTAAACAATTTGACTTCTCCAGTCACTGACGAGCAGCTGCTTTCATCATCTTCTTAA 3317
 QY 2665 ATCAGCCACCCATCGACACTTCCAAATTTGACCGGTGGATGAAGAAAGCCCTGGCAC 2724
 Db 3318 ATCAGACACCCCAACAGATACATCAAAATTTGATCTGTTGATCTCTGATAAATTTAGG 3377
 QY 2725 GAGGCGAG- --GGAGAGCGCCCAAGGCTGGGACAGCTGGCTTCCCTCCAGCAGCAG 2781
 Db 3378 GATGATAACGGAAGAAATTAATGACACTCTCAATGGATGGTATAAATAATGGAAG 3437
 QY 2782 CATCCAGAGCAGCCTTCTATGATGTTCCCTTCCGAGGTTCTTCGATGACACCGCTAT 2841
 Db 3438 CATCTGAACTGATCTATGAAATTTACCTTCCGAGGTTTTTTTGTGATGACATGGCTAC 3497
 QY 2842 CCCTTCCGGTGGCCGAAGCC 2861
 Db 3498 CCATATAATTTCCGAAGCC 3517

RESULT 6

US-09-442-100-1
 ; Sequence 1, Application US/09442100
 ; Patent No. 6359193
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Tian
 ; APPLICANT: Tao, Wufan
 ; APPLICANT: Wang, Weiyl
 ; APPLICANT: Zhang, Sheng
 ; APPLICANT: Yu, Wan
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
 ; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds

STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: PatentIn Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 FILING DATE: US/09/442,100
 CLASSIFICATION:
 PRIOR APPLICATION DATA:
 APPLICATION NUMBER: 08/411,111
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Misrock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 6523-003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 1:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 5720 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: cDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 1103..4402
 US-09-442-100-1

Query Match 17.98; Score 564.6; DB 4; Length 5720;
 Best Local Similarity 64.4%; Pred. No. 9.1e-139;
 Matches 979; Conservative 0; Mismatches 494; Indels 48; Gaps 7;

QY 1329 GGGCACTGATCTAGACGGAGTGACAGAGCCACAAAGGTGCGAAGGAGGAGACAAAGCTGG 1388
 DB 2869 GGGCCGGAATAGCTGGCGGCGAGCAACGGATCCACCGCGCACCGGCTCTCTCGTGAC 2928
 QY 1389 CAGAGACAAAAGCAGATTCAGACCTCCCGGTGCTCTCGCAGAGATAGC----- 1440
 DB 2929 CAGCTGCAAGAGATCAAGCAGCGCTCGCCCATCCCGAGGCGCAAGAGATCTCCAAGGA 2988
 QY 1441 -AGAGATGAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTATGCTTCAAAAT 1499
 DB 2989 GAAGGAGGAGGAGCGCAGAGATTCGGCATAGGAGTACTGCGCGCAAGCTTCAAGTT 3048
 QY 1500 CTTTCATGAGCAGCAGCTGGAGATGTATCAACAACTTACCAGCAGAGAGGTAGCGGAG 1559
 DB 3049 CTTTCATGAGCAGCAGATAGAGAACTGATCAAGTCTGATCGCCAGCAGCAGTATCGAA 3108
 QY 1560 GCTACAGCTGGAGAGGAATGCCAAGCTGGCTCTGTGAGGCGGAGCAGAGCAGAT 1619
 DB 3109 GAATCAGCTGGAGAGAGATGCAAAAGTGGGATGCGCCGATCAGCCCAATCCAGAT 3168
 QY 1620 GAGGAGATCTCTACCAAGAGAGTCTTAACATAACCGGCTGAAGAGGCGCAAGATGA 1679
 DB 3169 GAGGAAATGCTGAACCAAAAGAGAGCACTACATTCGATTGAAGCGCGCAAGATGA 3228
 QY 1680 CAGTCCATGTTGTGAAATCAAGACTTAGCATCGTCTTGGGAGTGGCT 1739
 DB 3229 CAAGAGCATGTTGCTCAAACTGAAGCCCATGGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3288
 QY 1740 CGCTTGAAGCTGA---CACTACAGCTCTGTACGCGCATGAAGACTCTCAGGAAGAAGGA 1796
 DB 3289 GGTGAGCAAAATCGATACCTCGNACCATTGTGATGCGATGAACACCTCGGGAAGCGGA 3348

QY 1797 TGTCTGACCGGAATCAAGTGGCCCATGTCAAGCTGAGAGGAGACATCTCTGGTGAAGC 1856
 DB 3349 CGTTCCTCAAGCGGAATCAGGTGGGCACAGCTGAGCCGAGAGGATATCTCTCGGGAAGC 3408
 QY 1857 AGACAATGAGTGGTGGTCAAACTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 1916
 DB 3409 CGACAATGAGTGGTGGTCAAACTCTACTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 3468
 QY 1917 TGTGATGAGTCAATACAGAGGCGGGATATGATGAGCTGTGATGATGATGATGATGATGATGAT 1976
 DB 3469 TGTGATGAGTCAATACAGAGGCGGGATATGATGAGCTGTGATGATGATGATGATGATGATGAT 3528
 QY 1977 CCCTGAGCAGCTGGCCGCTTCTACATTCAGAGTGTGACCTTGGCCATTTGAAGTGTCCA 2036
 DB 3529 CGAGGAGGAATGCGCAGATTCATATGCCGAGGTACCTGCGCGTGGACAGCGTTCA 3588
 QY 2037 CAAGATGGCTTTTATCCACCGGGACATCAAGCTTGACATCACTACTCATCTAGTGGATGG 2096
 DB 3589 CAAATGGCTTTTATCCACAGAGATCAAGCTTGACATCACTACTCATCTAGTGGATGG 3648
 QY 2097 TCATATTAGCTGACAGATTTGGCTCTGCACTGGATTTCAGTTCAGTTCACAAATTCCAA 2156
 DB 3649 ACACATAAAGCTCACCAGCTTTGGCTGTGACGGGATTCGATGACGACCACTCGAA 3708
 QY 2157 GTACTACC---AGAAAGGGAACACATGACAGAGCAGATGAGGAGCCCGGTGACCTCTG 2213
 DB 3709 GTACTACCAGGAGAACGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAGGAAT- 3767
 QY 2214 GGACGATGTTTCCAACTGTCGCTGTGGAGACAGTTTAAAGACCTCGAGCAGAGGCGCA 2273
 DB 3768 -----ACTCCGAGAACCGGACCGACCCACCGTCTGGAGGCGGAGAT 3813
 QY 2274 GAAGCAGCAGCAGAGGTGCTGCGCAGATCTCTTGTGGGACACCAAAATTCATCTGCTCC 2333
 DB 3814 GCGGATCACCAGAGATCTGCTGCCACCTGCTGGTGGGACCCCGCAACTACATAGCTCC 3873
 QY 2334 GGAGGTGCTTCTCCGAAAGGTACAGCAGCTGTGACTGTGTGAGGCTGCTGTGAT 2393
 DB 3874 CGAGGTGCTGAGAGGAGTGGGTACACGCGAGCTGTGCGACTGTGAGGAGTGGGCGGTAT 3933
 QY 2394 TCTTTTGAATGCTGTTGGGAGCGCGCTTTCTTGGCCCGCCACCCCGCCACAGAGAGCA 2453
 DB 3934 CTTTACGATGCTGTTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3993
 QY 2454 GCTGAAGGTGATCACTGGGAGAGCAGCTGATATCCCTTACGAGGTGAGGCTGAGCGC 2513
 DB 3994 ACAAAGGTATCACTGGGAGAGAAACGCTGATATTCGCGCCGAGCGCGAGTTATCCCG 4053
 QY 2514 TGAGGCGCGAGCTCATCAGGAGCTGTGCTGCGCGCTGACTGCGCGCTGGGAGGGA 2573
 DB 4054 CGAGGCTACGAGCTTGATAAGGAGGCTCTGTGCTGCGCTGACAAAGCGGTGGGCA---A 4110
 QY 2574 TGGGCGAGATGACCTCAAGGCACACCGCTTCTTCAACACCATCGACTTTTCCGCTGACAT 2633
 DB 4111 GAGCGTGGAGGAGTCAAGAGCCAGCAGCTTCTTCAAGGCGATCGACTTTG---GGACAT 4167
 QY 2634 CCGAAGCAGCTGCGCTTACGTCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2693
 DB 4168 GCGAAGCAGAGAGCGGCTTACATACCGGAATCAAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 4227
 QY 2694 TGACCGGTGATGAAGAAAGCCCTGGCAGAGGCGCGAGGAGAGCGGCAAGGCGCTG 2753
 DB 4228 TGATCCCGTGGATCCGGAGAGAGTGGCTCGAATGACTTCCAGCTAGCAGCGCGGATGA 4287
 QY 2754 GGACAGCTGCGCTCCCGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2813
 DB 4288 TGTGACCAAGTACCGCA-----CTTTCACGCGCTTTTTCGAATTTACCTT 4335
 QY 2814 CCGAGGTCTTTCGATGACAA 2834
 DB 4336 CCGTGCCTTCTTCGAGCAGAA 4356

RESULT 7

US-09-328-111-26/c
; Sequence: 26, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328.111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088.801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-328-111-26

Query Match 10.6%; Score 333; DB 4; Length 638;

Best Local Similarity 71.38; Pred. No. 3.4e-78;

Matches 454; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 1525 GTCATCAAAACCTACCAGCAGAAGTCCAGCGGAGGCTACAGCTGGAGCAGGAATGGCC 1584
DB 637 GTACTCAATCTCATCAGCAGCGCTACATCGTAAAAACAATAGAGAATGAATGATG 578
QY 1585 AAAGCTGGGCTCTGTGAGCGCCAGCAGGAGGAGATGAGGAAGATCCTCTACGAGAG 1644
DB 577 CGGTTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAGAA 518
QY 1645 TCTAATCAACCGGCTGAAGAGGCCAAGATGGAAGTCCATCTTTGTGAAATCAAG 1704
DB 517 TCTAATCACTCGGCTTAAAGGCTAAATGAGCAAGTCTATGTTGTGAAGATAAG 458
QY 1705 ACTTAGCATCGGTGCGCTTTGGGAGTGTGCGCTGCTGTAAGCTGGACACTCACGCT 1764
DB 457 ACACCTAGGATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATCTAAGCT 398
QY 1765 CTGTACGCCATGAAGACTCTCAGGAGAGGATGCTCTGACCGGAATCAAGTGGCCCAT 1824
DB 397 TTGTATGCAACAAAACCTCTCGAAATAAAGATGTTCTTCTCGAAATCAAGTGGCTCAT 338
QY 1825 GTCAGGCTGAGGGAGACATCTGTGCTGAAGCAGACATGAGTGGTGGTCTCAACTCTAC 1884
DB 337 GTTAAGCTGAGAGAGATATCTGCTGAGCTGAGCTGCAATGAATGAGTGGTCTGCTATAT 278
QY 1885 TACTCTCTCCAGGCAAGGAGCAGCGCTGTACTTTGTGAGTACTACATACAGGCGGGAT 1944
DB 277 TATTCATCCAAAGTAAGGCAATTTATATCTTTGTAATGGACTACATTTCTCTGGGGTGTAT 218
QY 1945 ATGATGACCTGTGATCAGATGAGGATCTTCCCTGAGCAGCTGGCCCGCTTCTACATTT 2004
DB 217 ATGATGACCTTATTAATAGGATGGCATCTTCCAGAAAGCTGGCCAGCATTTCTACATA 158
QY 2005 GCAGATTTGACCTGGCCATGAAAGTG- - -TCCACAGATGGGCTTTTATCCACGGGAC 2061
DB 157 GCAGACTTACTGTGAGTTGAAAGTGTCTTAAATGGGTTTATTTCATAGAGAT 98
QY 2062 ATCAAGCCTGACAACTACTCATCGACCTGGATGGTCTATATTAAGCTGACAGATTTTGGC 2121

DB 97 ATTAACCTGATAATATTTTATTGATCGTGATGCTATATAATTGACTGACTTTGGC 38
QY 2122 CTCTGCACTGGATTCAGTGGGACTCACAATTCCAAGT 2158
DB 37 CTCTGCACTGGCTTCAGATGGACACACGATCTTAAGT 1

RESULT 8

US-09-588-256-1
; Sequence 1, Application US/09588256
; Patent No. 6291665
; GENERAL INFORMATION:
; APPLICANT: Gaffney, Thomas
; APPLICANT: Flaviey, Albert
; APPLICANT: Gates, Krista
; APPLICANT: Wendland, Juergen
; APPLICANT: Ayad-Burieux, Yasmina
; APPLICANT: Dietrich, Fred
; APPLICANT: Philippsen, Peter
; TITLE OF INVENTION: Fungal Target Genes and Methods
; FILE REFERENCE: PB/5-30908A
; CURRENT APPLICATION NUMBER: US/09/588.256
; CURRENT FILING DATE: 2000-06-06
; NUMBER OF SEQ ID NOS: 24
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 1
; LENGTH: 2160
; TYPE: DNA
; ORGANISM: Ashbya gossypii
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (1)...(2160)
US-09-588-256-1

Query Match 8.2%; Score 258.8; DB 4; Length 2160;

Best Local Similarity 54.6%; Pred. No. 2e-58;

Matches 598; Conservative 0; Mismatches 447; Indels 51; Gaps 2;

QY 1637 AGAAGGAGCTTAATACAAACCGCTGAAGAGGGCCCAAGATGGACAGTCCATGTTGTGA 1696
DB 875 AAAAGGAGTCGAGTTCTCGCTTTGCGTAGGACACGCTATCCCTGGGAAGATTTCACA 934
QY 1697 AAATCAAGACTCTAGGCTCGGTGCTTTGGGGAAGTGTGCTGCTTTGAAGCTGGACA 1756
DB 935 CTGTAAAGTATAGGAAGGGTGCATTCGGTGGTCCGCTGCTGTCAGAGAAGAATA 994
QY 1757 CTCAGGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGATGTCTCTGAACCGGAATCAAG 1816
DB 995 CCGGTAATAATATACGCTATGAAGACATTGTTAAATATGAAATGTACAAGAAGATCAAT 1054
QY 1817 TGGCCCATGTCAAGGCTCAGAGGGACATCCCTGGCTGAAGCAGACATGAGTGGGTGGTCA 1876
DB 1055 TAGCCAGCTCAAGCGCAGAGGGATGTGTTGGCCGGAAGCGACTCTCCGTGGGTGGT 1114
QY 1877 AACTCTACTCTCTCCAGGCAAGGACAGCTCTACTTTGTGATGGACTACATACCAG 1936
DB 1115 GGTATATCTATCTTCCAGATGCCAGTACCTATATCTTGAATCAATGGAATTTTGGCCG 1174
QY 1937 CGGGGATATGATAGCCTGTGATCAGGATGGAGGTCTTCCCTGAGCAGCTGGCCCGCT 1996
DB 1175 GTGGTGACCTGATGACCATGTAATCAGGTGGCAGATATTCACCGAGGACGCTCACAGAT 1234
QY 1997 TCTACATTCAGAGTTGACCTGGCCATGAAAGTGTCCACAGATGGGCTTTATCCACC 2056
DB 1235 TCTACATCGCGAGTGTATCTGCGAATTTAGGCTTATACAAAGCTGGGCTTTATCCATA 1294
QY 2057 GGGACATCAAGCTGACAACTACTCATCGACCTGGATGGTGCATATTAACTGACAGATT 2116
DB 1295 GAGATATCAACCGGATTAACATCTGATGACATCAGGGGTCAATCAAACTTTCCGACT 1354
QY 2117 TTGGCTCTGCACTGGATTTCAGGTGGACTCAATTTCCAAGTACTACCAAGAGGAACC 2176

1355 TCGGTCTGTGACAGGGTTCCCAAAACGATGACTCCAACTACTACAAGAGCTGCTTC 1414
QY 2177 ACATGAGACAGGACAGCATGGAGCCCGGTGACCTCTGGAGCATGTTTCCAACTCTGCT 2236
Db 1415 AGGAGCAGGACAGCAGCAGAGAGCGGCGGAACATGGGCAATATCCGATCCGGTGGCG 1474
QY 2237 GTGGAGA-----CAGT 2248
Db 1475 GCGCAACGCGGCGGCAACAGAAACACCATGTTCTGCGAGCGCATCCACCTGACCATGA 1534
QY 2249 TAAAGACCTTGGAGAGGCGGCGAGCAGCAGCAGAGAGTGGCTTGGACATCTCTTG 2308
Db 1535 CARACAGCAGCAGATGCAAACTCGGCAAGTCCGCTAGGCTCATGGCTTACTCCACCG 1594
QY 2309 TCGGAGCACCNAATTACATCGCTCCGAGGTGCTTCTCCGAAAGGTACAGCAGCTCT 2368
Db 1595 TCGGTACGCCAGACTACATCGCCCGGAGATCTTCTCTACAGAGGCTACGGTTCAGGAGT 1654
QY 2369 GTGACTGTGGAGCGGTGGTGTGATTCTCTTTGAGATGCTGGTGGCAGCCGCTTCT 2428
Db 1655 CGACTGTGTCCTCGGCGCCATCATGTACGAGTGTCTGATCGGCTGGCCCGCTTCT 1714
QY 2429 TGGCCCCCACCACACAGAGCAGCAGCTGAGGTGATCACTGGGAGCAGCAGCTGCATA 2488
Db 1715 GCTGGAGACCCCGCAGAAACCTACAGGAAGATCATGAATTCGACAGCAGCAGCTGTGT 1774
QY 2489 TCCCTACGAGTGGTCTCAGCGCTGAGCGCCGAGACCTCATCACAAGCTGTGCTGG 2548
Db 1775 TCCAGAGCAGATCCACATCTCATAGGAGCAGGAGGACCTCATCCGCGGCTGCTCTCGC 1834
QY 2549 CGGCTGACTGGCGCTGGGCGAGGTGGGCGAGATGACCTCAAGGCGACACCGCTTCTCA 2608
Db 1835 AGCGCGAGCAGCGCTGCTGGCGATGCGGCGCAACAGAAATCAAAACCAACCCCTTCTTC 1894
QY 2609 ACACATCGACTTTTCCGCTGACATCCGAAAGCAGGCTGACCCCTAGCTCCCAACATCA 2668
Db 1895 GCGGCTGGAC---TGGGAGACCATCCGCCAGGTGCGGCGCTCCCTACATCCCAAGCTGT 1951
QY 2669 GCCACCCCATGGACAC 2684
Db 1952 CCAGCGCTCACCGACAC 1967

RESULT 9
US-08-878-989-11
Sequence 11, Application US/08878989
Patent No. 5885803
GENERAL INFORMATION:
APPLICANT: Bandman, Olga
APPLICANT: Hillman, Jennifer L.
APPLICANT: Corley, Neil C.
APPLICANT: Guebler, Karl G.
APPLICANT: Lal, Preeti
APPLICANT: Goli, Surya K.
APPLICANT: Shah, Purvi
TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
TITLE OF INVENTION: KINASES
NUMBER OF SEQUENCES: 21
CORRESPONDENCE ADDRESS:
ADDRESSEE: Incyte Pharmaceuticals, Inc.
STREET: 3174 Porter Drive
CITY: Palo Alto
STATE: CA
COUNTRY: USA
ZIP: 94304
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM compatible
OPERATING SYSTEM: DOS
SOFTWARE: FASTSEQ for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/878,989
FILING DATE:

CLASSIFICATION: 435
PRIOR APPLICATION DATA:
APPLICATION NUMBER:
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Billings, Lucy J.
REGISTRATION NUMBER: 36,749
REFERENCE/DOCKET NUMBER: PF-0321 US
TELECOMMUNICATION INFORMATION:
TELEPHONE: 415-855-0555
TELEFAX: 415-845-4166
TELEX:
INFORMATION FOR SEQ ID NO: 11:
SEQUENCE CHARACTERISTICS:
LENGTH: 1935 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
IMMEDIATE SOURCE:
LIBRARY: SYNORAT04
CLONE: 705365
US-08-878-989-11
Query Match 7.4%; Score 233.8; DB 2; Length 1935;
Best Local Similarity 51.7%; Pred. No. 7.4e-52;
Matches 609; Conservative 0; Mismatches 532; Indels 36; Gaps 2;
QY 1494 CAATTTCTTCATGGGACACACGCTGGAGAAATGTATCAAAACCTTACCAGCAGAGAGTTCAG 1553
Db 121 CAAGCTCACATTTGGAGAAATTTTATAGCAACCTAATTTTACAGCATGAGAGAGAGAAAC 180
QY 1554 CGGAGGCTACAGCTGGCAGGAAATGCCAAAGCTGGGCTCTGTAGGCGCCGAGCAGGA 1613
Db 181 CAGGCGAGAGAAATTAAGAAGTGGCCATGGAAGAAGAGATTAGCAGATGAAGAGAAAA 240
QY 1614 GCAGATGAGGAAGATCCTCTACAGAGAGAGTCTTAACACACCGCTTCAAGAGGCGCAA 1673
Db 241 GTTACGTCGATCACACACGCTCGCAAGAAACAGAGTTCTTACGGCTCAAAAGGACCG 300
QY 1674 GATGGACAGTCCATGTTTGTGAAATCAAGACTCTAGGCTCGGCTGCTTTGGGGAAGT 1733
Db 301 ACTTGGCTTGGATGACTTTGAGTCTCTGAAAGTTATAGAGAGGAGCTTTTGGAGAGT 360
QY 1734 GTGCTCGCTTTGAAGCTGGACACTCACGCTCTGTAGCCATGAAGACTCTCAGGNAGAA 1793
Db 361 CCGGTTGCTCCACAAAAGATACAGGCCATATCTATGCAATGAAGATATTGAGAAAGTC 420
QY 1794 GGATGCTCTGAACCGGAAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTCTGGCTGA 1853
Db 421 TGATATGCTTGAAGAGCAGGCTGGCCCATATCCGAGCAGAGAGAGATATTTTGGTGA 480
QY 1854 AGCAGACAATGAGTGGTGGTCAAACTCTACTACTCTCTCCAGCAGAGAGAGAGCTGTA 1913
Db 481 AGCAGATGGTGGCTGGTGGTGGTGGATGTTTACAGTTTTCAGCATAGAGAGAAATCTTA 540
QY 1914 CTTTGTGATGAGCTACATACAGCGGGGATATGATGAGCTCTGATCAGGATGAGAGT 1973
Db 541 TCTAATCATGGAATTTCTCCCTGGAGGTGATGATGATGATGATGATGATGATGATGATGAT 600
QY 1974 CTTCCCTCAGCACCTTGGCCGCTTCTACATTCAGAGTTGACCTTGCCCTGACCTGTA 2033
Db 601 CTTGACAG 660
QY 2034 CCACAAGATGGCTTTTATCCACCGGAGATCAAGCTTGCAGACATCTATCTATCTATCTAT 2093
Db 661 CCACCAAGTGGCTTTTATCCATCCGAGATATTAAGCAGAGAGAGAGAGAGAGAGAGAG 720
QY 2094 TGGTCATATTAGCTGACAGATTTTGGCTCTGCACTGGATTCAGGTTGAGTCAATTC 2153
Db 721 GGTGATGTAATTAATTTCTGATTTTGGTTTCTGATGATGATGATGATGATGATGATGAT 780
QY 2154 CAAGTACTACAG 2213

```

Db 781 TGAATTTTATAGAAATCTCACACACAAACCCACCA-----AGTGACTTCTCATTTTC 830
QY 2214 GGACGATGTTTCCAACTGTCGTGGAGACAGGTTAAAGACCTTGGAGCAGAGGCGCA 2273
Db 831 AGACATGAACCTCAAGAGAGAAAGCAGAACTTG-----864
QY 2274 GAAGCAGCAGAGAGGCTGCTGGCAGATCTCTTTCGGGACACCAAAATACATCGCTCC 2333
Db 865 GAAGAAGAACAGGAGACAACTGGCATATTCACAGTTGGGACACCAAGATTACATTCCTCC 924
QY 2334 GGAGTGCTTCTCCGCAAGAGGTACACGAGCTCTGTGACTGTTGGAGGCTGGTGTGAT 2393
Db 925 AGAAGTATTCATGCAGACTGGTTTACAAATTTGTGACTGTTTGGGAGTGAT 984
QY 2394 TCTCTTTGAGATGCTGGTGGCAGCGCTTCTTGGCCCCCACCACACAGAGACGA 2453
Db 985 TATGTATGAATGCTAATAGGATATCCACCTTTCGCTCTGAACACCTCAAGAAACATA 1044
QY 2454 GCTGAAGGTATCACTGGGAGAGCAGCTGCATATCCCTACGCGAGGTGAGCTCAGCGC 2513
Db 1045 CAGAAAAGTGATGAATGGAAGAAATCTGTGATTTCTCCAGAGGTACCTATATCTGA 1104
QY 2514 TGAGCGCCGAGACCTCATCAAGAGCTGTGCTGCGGGCTGACTGCGGCTGGGAGGA 2573
Db 1105 GAAAGCCAGGACTTAATTTCTCAGATTTTGTATTTCTGAAACAGAAATTTGGAATAG 1164
QY 2574 TGGGCGAGATGACCTCAAGGCACACCGCTTCTTCAACACCATCGACTTTTCCCGTGACAT 2633
Db 1165 TGGAGTAGAGAAATAAAGGTCATCCCTTTTGAAGGTGTCGACTGGGACCATATAG 1224
QY 2634 CCGAAGCAGGCTGCACCTAGTCCCTCCACCATCAGC 2670
Db 1225 GGAAGGCGCAGCAGCAATCCCTATAGAAATCAAAAGC 1261

```

```

RESULT 10
US-09-272-796-11
; Sequence 11, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PP-0321 US

```

```

; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAT04
; CLONE: 705365
US-09-272-796-11

```

```

Query Match 7.4%; Score 233.8; DB 4; Length 1935;
Best Local Similarity 51.7%; Pred. No. 7.4e-52;
Matches 609; Conservative 0; Mismatches 532; Indels 36; Gaps 2;
QY 1494 CAAATTTCTTCATGGAGCAACACGCTGGAGAAATGTCATCAAAACCTACCAGCAGAAGTCTAG 1553
Db 121 CAAGCTCAGATTTGGAGAAATTTTATAGCAACCTAATTTTACAGCATGAAGAGAGAGAAAC 180
QY 1554 CCGGAGGCTTACAGCTGGAGCAGGAAATGGCCAAAGCTTGGGCTCTGTGAGGCCGACGAGCA 1613
Db 181 CAGGCAGAGAAATTTAGAGTGGCCATGGAAGAAGAGGATTAGCAGATGAAGAGAGAAAC 240
QY 1614 GCAGATGAGGAAGATCTCTTACCAGAAAGAGTCTTAACATAACCGGCTGAAGAGGGCCAA 1673
Db 241 GTTACGTCGATCACAACACGCTCGCAAAAGAAACAGAGTTCTTTACGGCTCAAAAGACCCAG 300
QY 1674 GATGACAAAGTCCATGTTTGTGAAAATCAAGACTCTAGGCATCGTGGCTTTTGGGGAAGT 1733
Db 301 ACTTGGCTTGGATGACTTTGAGTCTCTGAAAGTTATAGGAGAGGAGCTTTTGGAGAGGT 360
QY 1734 GTGCTCGCTTGTAAAGCTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAA 1793
Db 361 GCGGTTGGTCCACAAAAAGATACAGGCCATATCTATGCAATGAAGATATTGAGAAAGT 420
QY 1794 GGATGCTCTGACCCGAGTCAAGTGGCCCATGTCAAGCTGAGAGGGAGCATCTCGGCTGA 1853
Db 421 TGATATGCTTGAAGAGAGCAGAGTGGCCCATATCCGAGCAGAAAGAGATATTTTGGTGA 480
QY 1854 AGCAGACAATGAGTGGTGGTCAAACTCTACTACTCTTCCAGGACAAAGCAGAGCTCTGA 1913
Db 481 AGCAGATGCTGCTGGTGGTCAAGATGTTTACAGTTTTCAGGATAGAGAGAAATCTTTA 540
QY 1914 CTTTGTGATGACTATACATACCAGCGGGGATATGATGACCTGTGATCAGGATGGAGGT 1973
Db 541 TCTAATCATGGAATTTCTCCCTGGAGGTGACATGATGACATTTGCTTAATGAAGAAAGAC 600
QY 1974 CTTCCCTGAGCACCTGGCCCGCTTCTACATTCAGAGTTGACCTTGGCCATTTGAAGTGT 2033
Db 601 CTTGACAGAGAGGAAACACACAGTTCTACATTTTCAGAGACTGTTTCTGGCAATAGATGCGAT 660
QY 2034 CCACAAAGATGGGCTTTATCCACCCGGACATCAAGCTGACACACTACTCATCGACTCGA 2093
Db 661 CCACAGTTGGTTTCATCCATCGGATATTAAAGCCACACACACTTTTATTTGGATGCCAA 720
QY 2094 TGGTCATATTAAGCTGACAGATTTTGGGCTCTGCACTCTGAGTTAGGTGAGCTACAATTC 2153
Db 721 GGGTCATGTAATAATTTATCTGATTTTGGTTCTGATGACGGGATTAAGAGAAAGCTCACAGGAC 780
QY 2154 CAAGTACTACCAAGAGGAAACCATGACAGCAGCAGCATGGAGCCCGGTGACCTCTG 2213
Db 781 TGAATTTTATAGAAATCTCACACACACCCACCA-----AGTGACTTCTCATTTTC 830
QY 2214 GGACGATGTTTCCAACTGTCGTGGAGACAGGTTAAAGACCTTGGAGCAGAGGCGCA 2273
Db 831 AGACATGAACCTCAAGAGAGAAAGCAGAACTTG-----864
QY 2274 GAAGCAGCAGAGAGGCTGCTGGCAGATCTCTTTCGGGACACCAAAATACATCGCTCC 2333

```

Db 865 GAAGAGAACAGGACCACTGGCATATTCACAGTTGGACACCAAGATACATTCCTCC 924
QY 2334 GGAGGTGCTTCTCCGAAAGGGTACAGCAGCTCTGTGACTGTGTGAGCGTGGTGTGAT 2393
Db 925 AAGATATTTCAGAGACTGGTTACAAACAAATGTGTGACTGTGTGAGGTGAT 984
QY 2394 TCTCTTTGAGATGCTGGTTGGGAGCGGCTTCTTTGGCCCCCACCACACAGAGCA 2453
Db 985 TATGTATGAATGCTAATAGGATATCCAGCTTCTCTCTGTAACACACCTCAAGAAACATA 1044
QY 2454 GCTCAAGGTGATCACTGGGAGAGCAGCTGCATATCCCTACCGAGTGGCTCAGCGC 2513
Db 1045 CAGAAAGTGAAGTGAAGAACTCTGTGATTTCTCCAGAGGTACCTATATCTGA 1104
QY 2514 TGAGGCCGAGACCTCATCACGAAGTGTGCTGCGCGGTGACTGCGCGCTGGCAGGGA 2573
Db 1105 GAAAGCCAAAGGACTTAATCTCAGATTTTGTATGTTCTGAACAGANTTGGAATAG 1164
QY 2574 TGCGGAGATGACCTCAAGGCACACCGTCTTCAACACCATTCGACTTTTCCCGTGACAT 2633
Db 1165 TCGAGTAGAAGAAATAAAGGTCATCCCTTTTGAAGGTGTGAGTGGGAGCACATAAG 1224
QY 2634 CCGAAGCAGGCTGCACCCCTAGCTGCCACCATCAGC 2670
Db 1225 GGAAGGCCAGCAGCAATCCCTATAGAAATCAAAAGC 1261

RESULT 11

US-08-860-150-6
; Sequence 6, Application US/08860150B
; Patent No. 5981205
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT APPLICATION NUMBER: US/08/860.150B
; EARLIER FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: PCT/EP95/05052
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: 94810746.1
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 3018
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (596)..(1990)
US-08-860-150-6

Query Match 7.1%; Score 225.2; DB 2; Length 3018;
Best Local Similarity 52.2%; Pred. No: 1.7e-49;
Matches 577; Conservative 0; Mismatches 493; Indels 36; Gaps 2;

QY 1565 AGCTGGAGCAGGAATGSCCAAGCTGGCTCTGTGAGGCCGAGCAGGACGATGAGGA 1624
Db 735 AGTTAGAAAGGTGATGGAAGAGAGGCGCTTAAAGATGAGGAGAACGACTCCGGAGAT 794
QY 1625 AGATCTCTACCAAGAGGAGTCTAACTACAAACCGGCTGAAGAGGCGCAAGATGACAAAGT 1684
Db 795 CAGCACATGCTCGGAAGGAACAGAGTTTCTGTTTGAAGAAACAGACTTGGATTGG 854
QY 1685 CCATGTTTGTGAATCAAGACTCTAGGCAATCGGTCCTTTGGGGAAGTGTGCTCGCTT 1744
Db 855 AGATTTTGAAGTCTTAAAGTAATAGGAGAGGAGCAATTTGTGAGGTACGGCTTGTTC 914
QY 1745 GTAAGTGGACACTCAGCTCTCTAGCCATGAAGACTCTCAGGAAGAGGATGCTCTGA 1804
Db 915 AGAAGAAAGATACGGGACATGTGTATGCAATGAAATCTCCGTAAGACAGATATGCTTG 974

QY 1805 ACCGGAATCAAGTGGCCCATGTCAAGCGTGTGAGAGGAGCATCTGCTGTGAAGCAGACAATG 1864
Db 975 AAAAAGAGCAGAGTTGGCCACATTCGTGGGAGCGGTGACATTCCTAGTGGAGCAGACAGTT 1034
QY 1865 AGTGGGTGCTCAAACTCTACTACTCTTCCAGGACAAGSAGAGCTGTACTTTGTGTG 1924
Db 1035 TGTGGTGTGNAATGTCTATAGTTTTCAGGATAAGCTAAACCTCTACCTAATCATG 1094
QY 1925 ACTATACACAGGCGGGGATATGATGACCTGTGATCAGGATGGAGTCTTCCCTCAGC 1984
Db 1095 AGTTCTCCCTGGAGGGACATGATGACCTGTGTTGATGAAAAAAGACACTCTGACAGAAG 1154
QY 1985 ACCTGGCCCGCTTCTACATTCAGAGTTGACCTTGGCCATTTGAAAGTGTCCACAAGATG 2044
Db 1155 AGGAGACTCAGTTTATATAGCAGAAAACAGTATTAGCCATAGACTCTATTCCCAACTG 1214
QY 2045 GCTTTATCCACCGGACATCAAGCTGACAACTACATCTGACCTGATGAGTGTGATGATTA 2104
Db 1215 GATTATCCACAGAGACATCAACACAGCAACCTCTTTTGGACAGCAAGGCCATGTGA 1274
QY 2105 AGCTGACAGATTTTGGCTCTGCACTGTGATTCAGTGGACTCAATTTCCAAGTACTACC 2164
Db 1275 AACTTTCTGACTTTGGTCTTTTGCAGAGACTGAAAAAAGACATAGGACAGAAATTTATA 1334
QY 2165 AGAAGGGAACCAACATGAGACAGGACAGCATGGAGCCCGTGCCTCTGGAGCATGTTT 2224
Db 1335 GGAATCTGAACCAACAGCTCC-----CCAGTGATTTTCACTTTCCAGAACATGAAT 1384
QY 2225 CCAACTCTGCTGTGAGACAGGTTTAAAGACCTTGAAGCAGAGGGCGCAGAAAGCAGCAC 2284
Db 1385 TCCAAAAGGAAAGCAGAAACCTGAAAAAGAAATAGCGTCAG-----1426
QY 2285 AGAGTCCCTGGCACATTCCTTGTGCGGACACCAATTTACATCGTCCGGAGGTGCTTC 2344
Db 1427 -----CTAGCCTTCTCCACAGTAGGCACTCTCTGACTACATTTGCTCTGAGGTGTTCA 1478
QY 2345 TCCGAAAGGTTACAGCAGCTCTGTGACTGTGGAGCGTGTGATTCCTTTTGTGAGA 2404
Db 1479 TGCAGACCGGTTACACAGCTCTGTGATTTGGTGTGCTTGGGGTGTATCATGTATGAGA 1538
QY 2405 TGTGTTGGGACCGGCTTTCTTGGCCCCCACCACAGAGAGCAGCTGAAGGTGA 2464
Db 1539 TGCTCATCGGCTACCCACCTTTCTGTCTGAGACCCCTCAAGAGACATATAAAGAGTGA 1598
QY 2465 TCAACTGGGAGACGCTGTCATATCCCTACGAGTGTGAGGCTCAGCCCTGAGGCCGAG 2524
Db 1599 TGAATCTGGAAGAACTTTGACTTTTCCCTCCAGAGTTCCCATCTCTGAGAAGCCAAAG 1658
QY 2525 ACCTCATCAGAGCTGTGCTGCGCGGTGACTGCGCGCTGGGAGGATGGGAGATG 2584
Db 1659 ATCTAATTTTGAAGTTCTGCTGTGAATGGAACATAGAAATTTGAGCTCTCTGAGGTGAGG 1718
QY 2585 ACCTCAAGCAGACCGTTCTTCAACACCATCGACTTTTCCCGTGTGACATCCGAAAGCAGG 2644
Db 1719 AATAAAGTAACTCTTTTGTGAAGCGTTGACTTGGGAACATATCAGAGAGAGACCTG 1778
QY 2645 CTGCACTTACCTCCCAACCATCAGC 2670
Db 1779 CTGCAATATCTATTGAAATCAAAAGC 1804

RESULT 12

US-09-338-132-6
; Sequence 6, Application US/09338132
; Patent No. 6040164
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT APPLICATION NUMBER: US/09/338.132
; CURRENT FILING DATE: 1999-06-22
; EARLIER APPLICATION NUMBER: 08/860.150

EARLIER FILING DATE: 1997-06-19
 EARLIER APPLICATION NUMBER: PCT/EP95/05052
 EARLIER FILING DATE: 1995-12-20
 EARLIER APPLICATION NUMBER: 94810746.1
 EARLIER FILING DATE: 1994-12-22
 NUMBER OF SEQ ID NOS: 18
 SOFTWARE: PatentIn Ver. 2.0
 SEQ ID NO 6
 LENGTH: 3018
 TYPE: DNA
 ORGANISM: Homo sapiens.
 FEATURE:
 NAME/KEY: CDS
 LOCATION: (596)..(1990)
 US-09-338-132-6

Query Match 7.1%; Score 225.2; DB 3; Length 3018;
 Best Local Similarity 52.2%; Pred. No. 1.7e-49;
 Matches 577; Conservative 0; Mismatches 493; Indels 36; Gaps 2;

QY 1565 AGCTGGACGAGAAATGCCAAAGCTGGCTCTGTGAGCGCGAGCAGGAGCAGATGAGGA 1624
 DB 735 AGTTAGAAAGGTGATGAGAGAGAGAGGCGCTAAAGATGAGGAGAAACGACTCCGGAGAT 794

QY 1625 AGATCCTCTACCAAGAGGAGTCTAACTACAAACCGCTGAAGAGCGGCCCAAGATGGACAAGT 1684
 DB 795 CAGCACATGCTCGGAAGGAACAGAGTTCTTCGTTTGAAGAGAACAGACTTGGATTGG 854

QY 1685 CATGTTTGTGAATCAAGACTCTAGGACATCGGTGCTTGGGGAAGTGTGCCCTGCTT 1744
 DB 855 AAGATTTTGTAGTCCCTTAAAGAGTAATPAGGCAGAGGAGCATTTTGGTGAGGTACGCTTGTTC 914

QY 1745 GTAAGCTGAGCACTCACGCTCTGTAGCGCATGAGCACTCTCAGGAAGAGGATGCTCTGA 1804
 DB 915 AGAAGAAAGATACGGGACATGTGTGCAATGAATAACTCCGTAAAGACAGATGCTTGG 974

QY 1805 ACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGAGCATCTCTGCTCAAGCAGACAATG 1864
 DB 975 AAAAGAGCAGGTTGGCCACATTCGTGCGGAGCGTGACATCTAGTGGAGGCGAGACGTT 1034

QY 1865 AGTGGTGTCAAACTCTACTACTCTCTCCAGGACAGACGCTGTACITTTGTGATGG 1924
 DB 1035 TGTGGGTTGTGAATGTTCTATAGTTTTCAGGATTAAGCTAAACCTCTACCTTAATCATGG 1094

QY 1925 ACTACATACCAAGCGGATGATGATGAGCTCTGCTCAGGATGAGGTCTTCCCTGAGC 1984
 DB 1095 AGTTCTGCTGGAGGAGCATGATGACCTGTTGATGAAGAAACACACTCTGACAGAG 1154

QY 1985 ACCTGGCCCGTCTTACATTCAGAGTTTGACCCCTGGCCATTGAAAGTGTCCACAAGATGG 2044
 DB 1155 AGGAGACTCAGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATTACCAACTTG 1214

QY 2045 GCTTTATCCCGGAGACATCAAGCTGTACACATACTCATCGACCTGGATGTCATATTA 2104
 DB 1215 GATTTCATCCAGACACATCAACACAGAACCTTCTTTTGGACAGCAAGGCGCCATGTA 1274

QY 2105 AGCTGACAGATTTTGGCTCTGCACTGAGTTGAGTTCAGCTGAGCTCACAAATCCAGTACTACC 2164
 DB 1275 AACTTCTGACTTTGGCTTTGACAGGACTGAAAAAGACATAGGACAGAAATTTATA 1334

QY 2165 AGAAGGGAACACATGAGACAGGAGCAGCATGGAGCCCGGTGACCTCTGGGACGATGTT 2224
 DB 1335 GGAATCTGAACACAGCCTCC-----CCAGTGATTTCACITTTCCAGAACATGAAT 1384

QY 2225 CCAACTGTGCTGTGGAGACAGGTTAAAGACCTCTGGACAGAGGCGGAGAGAGCAGCACC 2284
 DB 1385 TCCAAAGGAAGACAGAAACCTGGAAAAAGAAATAGACGTCAG----- 1426

QY 2285 AGAGTGTGCTGACATTTCTTTGTCGGGACACCAAAATTACATCTCGTGGAGGTGCTTC 2344
 DB 1427 -----CTAGCTTCTCCACAGTAGGACATCTCTGACTACATCTCTGAGGTGTTCA 1478

QY 2345 TCCGAAAGGAGTACACGAGCTCTGTGACTGTGTGAGGCGTGGTGTGATTTCTTTTGAGA 2404

DB 1479 TGCACACCGGGTACAAAGCTCTGTGATTTGGTGTGGGTGATCATGTATGAGA 1538
 QY 2405 TGCTGTTGGGAGCGGCTTCTTGGGCCCCACCCACACAGAGAGGAGCTGAAGGTGA 2464
 DB 1539 TGCTCATCGGCTACCCAGCTTTCTGTCTGAGACCCCTCAAGAGACATATAAAGAGGTGA 1598

QY 2465 TCAACTGGGAGAGCACGCTGTCATATCCCTACGACGCTGAGGCTCAGCGCTGAGGCCCGAG 2524
 DB 1599 TGAAGTGAAGAAACTTTTGACTTTTCTCCAGAAGTTCCCACTCTCTGAGAAGCCAGG 1658

QY 2525 ACCTCATCACGAAGCTGTCTGCGGCTGACTGCGGCTGGGAGGATGGGAGATG 2584
 DB 1659 ATCTAATTTGAGGTTCTGCTGTAATGGGAACATAGAATTTGAGCTCTCTGAGTTGAGG 1718

QY 2585 ACCTCAAGCAGACCGGCTTCTTCAACACCATGCACTTTTCCGCTGACATCCCAAGACAGG 2644
 DB 1719 AAATAAAAGTAACCTCTTTTGAAGGCGTTGACTGGGAACATATCAGAGAGACCTG 1778

QY 2645 CTGACCCCTACGTCCTCCCAACATCAGC 2670
 DB 1779 CTGCAATATCTATTGAATCAAAAGC 1804

RESULT 13
 US-08-860-150-1
 ; Sequence 1, Application US/088601508
 ; Patent No. 5981205
 ; GENERAL INFORMATION:
 ; APPLICANT: Hemmings, Brian A.
 ; APPLICANT: Millward, Thomas A.
 ; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
 ; FILE REFERENCE: 4-20265/A/PCT
 ; CURRENT APPLICATION NUMBER: US/08/860,150B
 ; EARLIER FILING DATE: 1997-06-19
 ; EARLIER APPLICATION NUMBER: PCT/EP95/05052
 ; EARLIER FILING DATE: 1995-12-20
 ; EARLIER APPLICATION NUMBER: 94810746.1
 ; EARLIER FILING DATE: 1994-12-22
 ; NUMBER OF SEQ ID NOS: 18
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 1
 ; LENGTH: 2101
 ; TYPE: DNA
 ; ORGANISM: Drosophila melanogaster
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: (132)..(1499)
 ; US-08-860-150-1

Query Match 7.1%; Score 222.6; DB 2; Length 2101;
 Best Local Similarity 52.5%; Pred. No. 6.9e-49;
 Matches 557; Conservative 0; Mismatches 474; Indels 30; Gaps 2;

QY 1563 ACAGCTGGAGCAGGAAATGGCCAAAGCTGGCTCTGTGAGCGCGAGCAGGAGCAGATGAG 1622
 DB 281 AAAGCTGGAGGCTCAGCTCAAGGACGAGAGCTTGTGCGAGGCGCAGCGCCAGGAGG 340

QY 1623 GAAGATCTCTTACCAGAGGAGTCTAATCAACCGGTGAAGAGGCGGCAAGATGAGCAA 1682
 DB 341 TCTGACGATGTCGCAAGAGGAGCAGGATATCTCCGGTGAAGCGATTTCGCGCTCGGTGT 400

QY 1683 GTCATGTTTGGAAAAATCAAGACTCTAGGCATCGTGGGAGTGTGGGAGTGTGCTCGC 1742
 DB 401 GGAGACTTTGAGGCCCTCAAAAGTCAATCGAGCGCGCGCTTCGGTGAAGTTCGTTTGT 460

QY 1743 TTGTAAGCTGACACTCAGCTCTCTAGCCCATGAAGACTCTCAGGAAGAGATGTCTCT 1802
 DB 461 GCAGAAAAAGGACACTGGACATGTCTGCCCATGAAGTGTCTGCCAAAGCGACATGCT 520

QY 1803 GAACCGGAATCAAGTGGGCCCATGTCAAGCTGAGAG---GGACATCTGGCTGAAGCAGA 1859
 DB 521 GGAAGAGGAGGAGGTGGACACAGTACGGGCCGAGGCTGTGATGTCTTCTGTCGAGGCCGA 580

Fri Jan 17 11:17:04 2003

```

; SEQ ID NO 1
; LENGTH: 2101
; TYPE: DNA
; ORGANISM: Drosophila melanogaster
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (132)..(1499)
; US-09-338-132-1

Query Match      7.1%; Score 222.6; DB 3; Length 2101;
Best Local Similarity 52.5%; Pred: No. 6.9e-49;
Matches 557; Conservative 0; Mismatches 474; Indels 30; Gaps 2;

1563 ACACCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGAGCGCGGACGAGCAGATGAG 1622
Db 1563 ACACCTGGAGCAGGCTGAGGAGCAGAGAGCTTGTGCGAGGCGCAGCGCAGGAGCG 340
1623 GAAGATCCTCTACAGAGGAGTCTAATCAACCGCTGAAGAGGCGCAAGATGACAA 1682
Db 1623 GAAGATCCTCTACAGAGGAGTCTAATCAACCGCTGAAGAGGCGCAAGATGACAA 1682
341 TCTGAGCATGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
Db 341 TCTGAGCATGCCAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 400
1683 GTCCATGTTTGTGAAATCAAGACTTAGGCACTCGGTGCGCTTTGGGGAAGTGTCTCG 1742
Db 1683 GTCCATGTTTGTGAAATCAAGACTTAGGCACTCGGTGCGCTTTGGGGAAGTGTCTCG 1742
401 GGAGAGCTTTGAGGCGCTCAAGTCAATCGGACGCGCGCTTCGGTGAAGTGTGGT 450
Db 401 GGAGAGCTTTGAGGCGCTCAAGTCAATCGGACGCGCGCTTCGGTGAAGTGTGGT 450
1743 TTGTAAGCTGGACACTACGCTCTGTAGCCATGACGCTGAGAG---GGACATCTCGTGGTGAAGCAGA 1859
Db 1743 TTGTAAGCTGGACACTACGCTCTGTAGCCATGACGCTGAGAG---GGACATCTCGTGGTGAAGCAGA 1859
1803 GAACCGGATCAAGTGGCGCATGTCAAGCTGAGAG---GGACATCTCGTGGTGAAGCAGA 580
Db 1803 GAACCGGATCAAGTGGCGCATGTCAAGCTGAGAG---GGACATCTCGTGGTGAAGCAGA 580
521 GCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1919
Db 521 GCAAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1919
1860 CAATGAGTGGGTGGAATAAATCTACTACTCTTCCAGGAGGAGGAGGAGGAGGAGGAG 640
Db 1860 CAATGAGTGGGTGGAATAAATCTACTACTCTTCCAGGAGGAGGAGGAGGAGGAGGAG 640
581 TCATCAGTGGGTGGAATAAATCTACTACTCTTCCAGGAGGAGGAGGAGGAGGAGGAG 1979
Db 581 TCATCAGTGGGTGGAATAAATCTACTACTCTTCCAGGAGGAGGAGGAGGAGGAGGAG 1979
1920 GATGAGTACATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 700
Db 1920 GATGAGTACATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 700
641 AATGAGGAGTCTTCCGCTGGGTGATGATGATGATGATGATGATGATGATGATGATGAT 2039
Db 641 AATGAGGAGTCTTCCGCTGGGTGATGATGATGATGATGATGATGATGATGATGATGAT 2039
1980 TGAGACCTGGCGCGTCTTACATTTCCAGATTTGACCTTGGCCATTTGAAAGTGTCCACAA 760
Db 1980 TGAGACCTGGCGCGTCTTACATTTCCAGATTTGACCTTGGCCATTTGAAAGTGTCCACAA 760
701 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2099
Db 701 CGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2099
2040 GATGGGCTTTATCCACCGGAGCATCAAGCTGACATCACTACTACTACTACTACTACTACT 820
Db 2040 GATGGGCTTTATCCACCGGAGCATCAAGCTGACATCACTACTACTACTACTACTACTACT 820
761 ACTCGGTTTATACACAGGAGGATATCAAGCCGATATCAAGCCGATATCAAGCCGATAT 2159
Db 761 ACTCGGTTTATACACAGGAGGATATCAAGCCGATATCAAGCCGATATCAAGCCGATAT 2159
2100 TATTAAGCTGACAGATTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
Db 2100 TATTAAGCTGACAGATTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 880
821 TCTGAAGCTCTCCGACTTCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
Db 821 TCTGAAGCTCTCCGACTTCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2219
2160 CTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 940
Db 2160 CTACAGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 940
881 TTATCGGAGCTTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2279
Db 881 TTATCGGAGCTTGTGCGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2279
2220 TGTTCCTCAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 973
Db 2220 TGTTCCTCAACTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 973
941 GGACTCCAAAGCAGCTGCGG-----AGTCTGGAAGCG 2339
Db 941 GGACTCCAAAGCAGCTGCGG-----AGTCTGGAAGCG 2339
2280 GCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033
Db 2280 GCACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1033
974 AAATCGAGCGGCGCTCGCTACGACCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2399
Db 974 AAATCGAGCGGCGCTCGCTACGACCGCTGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2399
2340 GCTTCTCCCAAGGAGTACCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
Db 2340 GCTTCTCCCAAGGAGTACCGAGCTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1093
1034 ATTTCTGACAGCTGGCTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2459
Db 1034 ATTTCTGACAGCTGGCTACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2459
2400 TGAGATGCTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1153
Db 2400 TGAGATGCTGGTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1153
CGAAGTCTGATGGCTATCTCTCATTTCTGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 2519
Db CGAAGTCTGATGGCTATCTCTCATTTCTGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 2519
GGTGATCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1213
Db GGTGATCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1213
GGTGATCACTGGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2579
Db CCGAGAGCTCATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 2579
2520 CCGAGAGCTCATACGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1273
Db 1214 CAAGGAGAGGATCATCAACTTCTGCTGCGAGGAGGAGGAGGAGGAGGAGGAGGAG 2620
QY 2580 AGATGACCTCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1314
Db 1274 TGGAGGATCTGAAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1314

RESULT 14
US-09-338-132-1
; Sequence 1, Application US/09338132
; Patent No. 6040164
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT APPLICATION NUMBER: US/09/338,132
; EARLIER FILING DATE: 1999-06-22
; EARLIER FILING DATE: 1997-06-19
; EARLIER FILING DATE: 1995-12-20
; EARLIER FILING DATE: 1994-12-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0

```


QY 2460 GTGATCAACTGGAGAGACGCTGCATATCCCTACGAGGTGAGGCTCAGCGCTGAGGC 2519
 Db 1154 GTGATGAATCGCGGAGACGCTGATATTTCCCCCAAGAGATCCCATATCGAGGAGGC 1213
 QY 2520 CCGAGACCTATCAACGAAGCTGTGTCGCGGCTGACTCCGCGCTGGGAGGATGGGCG 2579
 Db 1214 CAAGGAGAGCATCATCAACTTCTGTGCGAGGCGGATCGCGCTGGTTCCAGCGTCGTC 1273
 QY 2580 AGATGACCTCAAGGCACACCCGCTTCTCAACACCATCGACT 2620
 Db 1274 TGAGAGATCTGAAGTCGTCGCTTCTTCCGGGAGTTGACT 1314

RESULT 15

US-09-328-111-66
 ; Sequence 66, Application US/09328111
 ; Patent No. 6262333

GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astle, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/328,111
 ; CURRENT FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: US 60/088,801
 ; EARLIER FILING DATE: 1998-06-10
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: FastSeq for Windows version 3.0
 ; SEQ ID NO 66

; LENGTH: 678

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc_feature

; LOCATION: (1)...(678)

; OTHER INFORMATION: n = A,T,C or G

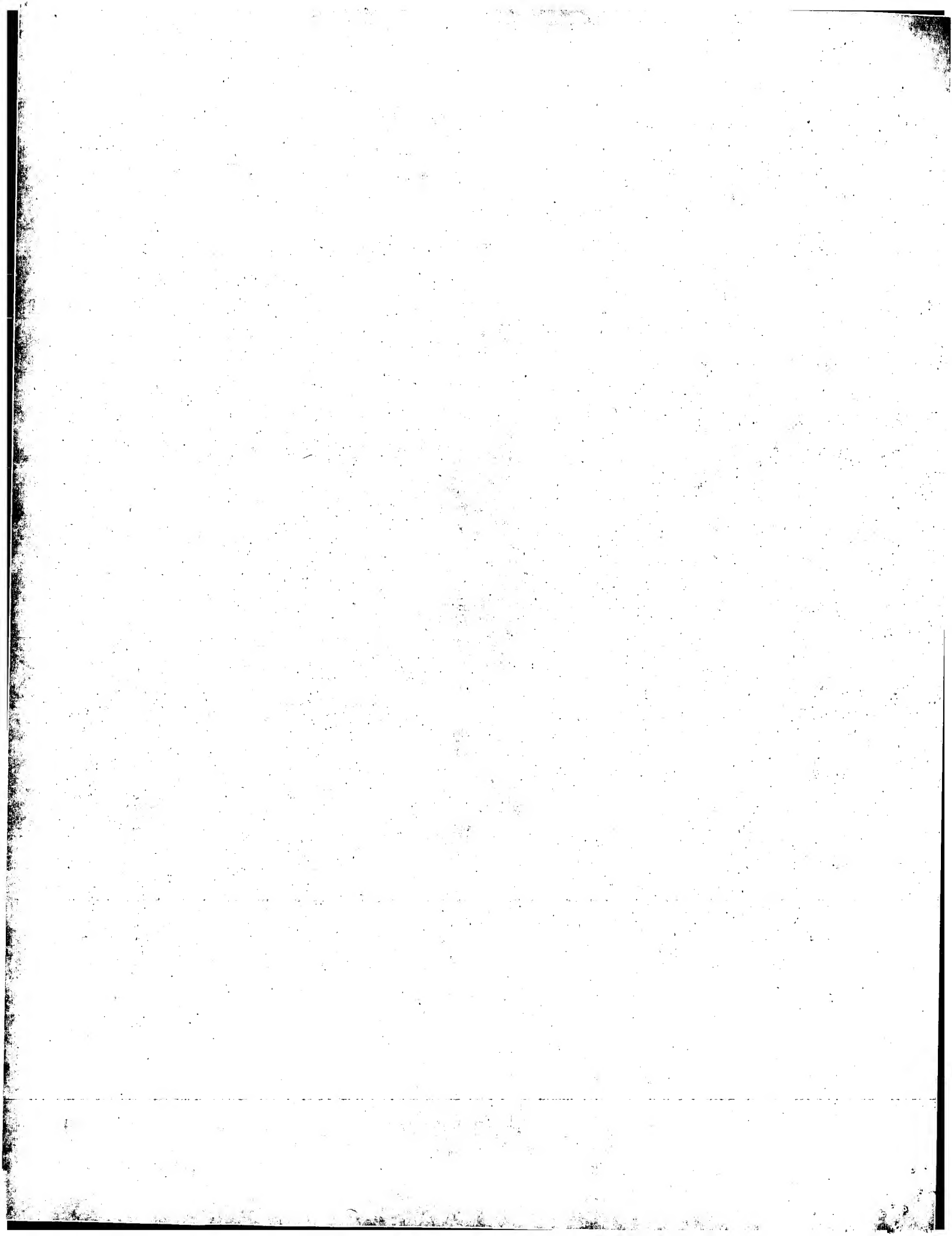
US-09-328-111-66

Query Match 6.8%; Score 214.4; DB 4; Length 678;
 Best Local Similarity 68.8%; Pred. No. 6e-47; Mismatches 138; Indels 2; Gaps 1;
 Matches 308; Conservative 0;

QY 1529 TCAAAACCTTACAGCAGAGAGGTTCAGCGGAGGCTACAGCTGGAGCAGGAAATGGCCAAAG 1588
 Db 3 TCAAAATCTCATCAGCAGCGCTTACATCGTAAAAACAATTAGAGATGAATGATCCGGG 62
 QY 1589 CTGGGCTCTGTGAGGCGGAGGAGCAGATGAGGAAGATCTCTACCAAGAGGAGTCTA 1648
 Db 63 TTGGATTATCTCAAGATGCCAGGATCAAAATGAGAAAGATGCTTTGCCAAAAAAGATCTA 122
 QY 1649 ACTACAACCGGCTGAGAGGCGCAGATGGACAAGTCCATGTTTGTGAAAAATCAAGACTC 1708
 Db 123 ATTACATCCGCTTAAAGGGGCTAAATGGACAAGTCTATGTTGTGAAGATAAAGACAC 182
 QY 1709 TAGGCATCGGTCTTTGGGGAAGTGTGCTCGCTTGTAAAGCTGGACACTCAGCTCTGT 1768
 Db 183 TAGGAATAGGACATTTGTGAGTCTGTCTAGCAAGAAAAAGTAGATACTAAGGCTTGT 242
 QY 1769 ACGCCATGAAGACTCTCAGGAAGAAGAGTGTCTTGAACCGGAATCAAGTGGCCCATGTCA 1828
 Db 243 ATGCAAGAAAAAATCTTCCGAAAGAAGATGTTCTTCTTCCGAATCAAGTCGCTCATGTTA 302

QY 1829 AGGCTGAGAGGAGACATCCCTGGCTGAAGCAGACAATGAGTGGTGGTCAAACTCTACT 1888
 Db 303 AGGCTGAGAGAGATATCTCTGGCTGAGCTGACAATGAATGGTAGTTCGCTCTATATT 362
 QY 1889 CTTCCAGGACAAGGACAGCCTGTACTTTGTGATGGACTACATACAGCGGGGATATGA 1948
 Db 363 CATTCCAAGATAAGGCG--CATTTATCCTTGTAAATGGCTACATTCCTNGGGGTGATATGA 420
 QY 1949 TGAGCCTGCTGATCAGGATGGAGGCTTT 1976
 Db 421 AGAGCCCATTAATANAATGGCATCTT 448

Search completed: January 16, 2003, 21:57:26
 Job time : 134.191 secs



Result No.	Score	Query Match	Length	DB ID	Description	
1	1035.2	32.8	2043	9	US-09-836-392-2	
2	546.6	17.3	676	9	US-09-764-868-214	
C	3	333	10.6	638	10	US-09-879-536-26
	4	231.2	7.3	1689	9	US-09-938-842A-1861
	5	228	7.2	1452	9	US-09-938-842A-2402
	6	225.2	7.1	3583	9	US-09-974-298-152
7	214.4	6.8	678	10	US-09-879-536-66	
8	192.2	6.1	734	9	US-09-764-868-196	
9	169.4	5.4	1818	10	US-09-771-161A-89	
10	161.8	5.1	2549	10	US-09-880-107-3691	
11	156.6	4.9	1635	10	US-09-880-107-2340	
12	149.8	4.7	3407	10	US-09-871-845-1	
13	142	4.5	1398	9	US-09-938-842A-633	
14	138.8	4.4	1416	9	US-09-938-842A-2503	
15	135.6	4.3	1735	9	US-09-764-868-58	
16	134.8	4.3	1244	10	US-09-771-161A-38	
17	134.8	4.3	1393	10	US-09-771-161A-37	
18	133.6	4.2	1257	10	US-09-799-875-15	
19	133.6	4.2	1826	10	US-09-799-875-13	
20	133.6	4.2	1826	10	US-09-799-875-13	

0 1 2 3 4 5 6 7 8 9 A B C D E F

QY 1765 CTGTACGCCATGAAGACTCTCAGGAGAGAGGATCTCTGAACCGGATCAAGTGGCCAT 1824
 Db 395 CTGTACGCCATGAAGACTCTCAGGAGAGAGGATCTCTGAACCGGATCAAGTGGCCAT 454
 QY 1825 GTCAAGGCTCAGAGGACATCCCTGGGTGAAGACAGCAATGAGTGGGTGTCAAACTCTAC 1884
 Db 455 GTCAAGGCTCAGAGGACATCCCTGGGTGAAGACAGCAATGAGTGGGTGTCAAACTCTAC 514
 QY 1885 TACTCTCTCCAGGACAGGACGCTGTACTTTGTGATGAGTACATACACGCGGGAT 1944
 Db 515 TACTCTCTCCAGGACAGGACGCTGTACTTTGTGATGAGTACATACACGCGGGAT 574
 QY 1945 ATGATGAGCCTGTGATCAGGATGGAGGCTTCCCTGAGCAGCTGCCGCTCTCTACAT 2004
 Db 575 ATGATGAGCCTGTGATCAGGATGGAGGCTTCCCTGAGCAGCTGCCGCTCTCTACAT 634
 QY 2005 GCAGAGTTGACCTGGCCATTGAAGTGTCCACAAGATGGGCTTTATCCACCGGGATC 2064
 Db 635 GCAGAGTTGACCTGGCCATTGAAGTGTCCACAAGATGGGCTTTATCCACCGGGATC 694
 QY 2065 AAGCCTGACACATPACATCTGAGCTGGATGGTGTCAATTAAGCTGACAGATTTTGGCTC 2124
 Db 695 AAGCCTGACACATPACATCTGAGTGGTGTCAATTAAGCTGACAGATTTTGGCTC 754
 QY 2125 TGCATGGATTCAGTGGACTCAAAATTCACAGTACTACAGAAAGGAAACACATGAGA 2184
 Db 755 TGCATGGATTCAGTGGACTCAAAATTCACAGTACTACAGAAAGGAAACACATGAGA 814
 QY 2185 CAGGACAGTATGAGCGGCTGACCTCTGGAGAGATTTTCCAACTGCTGCTGGAGAC 2244
 Db 815 CAGGACAGTATGAGCGGCTGAGGACCTCTGGAGATTTTCCAACTGCTGCTGGAGAC 874
 QY 2245 AGGTTAAGACCTTGGAGAGGCGCAGAGCAGCAACAGAGTGCCTGGCAGATTC 2304
 Db 875 AGGTTAAGACCTTGGAGAGGCGCAGAGCAGCAACAGAGTGCCTGGCAGATTC 934
 QY 2305 CTGTGCGGACCAAAATACATCGCTCGGAGGTGCTTCCCAAGAGGTACACGAC 2364
 Db 935 CTGTGCGGACCTCAAAATACATCGCAGCGAGGTGCTTCCCAAGAGGTACACGAC 994
 QY 2365 CTGTGAGTGGTGGAGGCTCGGTGATTTCTTTGAGATGCTGGTGGCAGCGCT 2424
 Db 995 CTGTGAGTGGTGGAGGCTCGGTGATTTCTTTGAGATGCTGGTGGCAGCGCT 1054
 QY 2425 TCTTGGCCCCCCCCACACAGACGAGCTGAAGTGTCAACTGGGAGACGACGCTG 2484
 Db 1055 TCTTGGCACCTTCTCCACAGAACCCAGCTGAAGTGTCAACTGGGAGACGACGCTG 1114
 QY 2485 CATATCCCTAGCAGGTGAGGCTCAGCGCTGAGGCGGAGACCTCATCAGAGCTGTGC 2544
 Db 1115 CATATCCAGCCAGGTGAAGTGTGAGCCTGAGGCGGAGGACCTCATCAGAGCTGTGC 1174
 QY 2545 TCGCGGCTGACTGCGGCTGGCAGGATGGGCGAGATGACCTCAAGGACACCCGTC 2604
 Db 1175 TCGTCCGAGACACCGCTGGGCGGAAATGGGCGGATGACCTCAAGGCGGACCCCTTC 1234
 QY 2605 TTCAACACCATCGACTTTTCCGTCGATCCGAAGAGAGGCTGACCTAGCTCCCGACC 2664
 Db 1235 TTCAAGCGCATTTCTTCCAGTACATCCGAAGAGAGGACGACCCCTAGCTTCCCGACC 1294
 QY 2665 ATCAGCCACCCCATGACACCTTCCAAATTTTACCGGTTGATGAAGAAGCCCTGGGAC 2724
 Db 1295 ATCAGCCACCCCATGACACCTTCCAAATTTTACCGGTTGATGAAGAAGCCCTTGGAC 1354
 QY 2725 GAGGCGGCGGAGAGGCGGACGCTGGGACGCTGGCTCCCGGACGAGCAAGCAT 2784
 Db 1355 GATGCGGCGGAGGAGGCGGACGCTGGGACGCTGGGACGCTGGGACGCTGGGACGAT 1414
 QY 2785 CCAGAGCAGCGCTTCTATGAGTTTACCTTCCGAGGTTTCTGATGACACAGGCTATCCC 2844
 Db 1415 CCTGAGCAGCGATTTTACGAATTCACCTTCCGAGGTTTCTGATGACAAATGGCTACCCC 1474

QY 2845 TTCCGGTCCCGAAGCCCTCAG-----AGCCCGCAGAGAGTGCAGACCCAGGG 2892
 Db 1475 TTTCGATCCCAAGAGCTTTCAGAGCAGAGAGCTTTCAGAGCTCAGATTTAGAA 1534
 QY 2893 GATGGCGACTTGAAGTGGCGGCGGAGGCTGCCAGCGGTGTAGCTGTA 2942
 Db 1535 AGCTCTGATCTGTGGTGCAGACTGAAGCTGCCAGCCTGTGTACGTGA 1584
 RESULT 2
 US-09-764-868-214
 ; Sequence 214, Application US/09764868
 ; Patent No. US20020168711A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Rosen et al.
 ; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
 ; FILE REFERENCE: PT232
 ; CURRENT APPLICATION NUMBER: US/09/764,868
 ; CURRENT FILING DATE: 2001-01-17
 ; Prior application data removed - refer to PALM or file wrapper
 ; NUMBER OF SEQ ID NOS: 1510
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 214
 ; LENGTH: 676
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: SITE
 ; LOCATION: (628)
 ; OTHER INFORMATION: n equals a,t,g, or c
 US-09-764-868-214
 Query Match 17.3%; Score 546.6; DB 9; Length 676;
 Best Local Similarity 89.5%; Pred. No. 1.8e-138;
 Matches 607; Conservative 2; Mismatches 67; Indels 2; Gaps 2;
 QY 1492 TTCAAAATTTCTGATGAGCAACACGTTGGAGATGTCTCAAAACCTTACCAGCAGAGGTC 1551
 Db 1 TTTAAGTTCTTCATGAGGAGCAGCAGCTGGAGATGTCTCAAAACCTTACCAGCAGAGGTT 60
 QY 1552 AGCGGAGGCTACAGCTGGAGAGGAATGGCCAAAGCTGGGCTCTGTAGGCGGAGCAG 1611
 Db 61 AACCGGAGGCTGCAGCTGGAGAGGAATGGCCAAAGCTGGGCTCTGTAGGCTGAGCAG 120
 QY 1612 GAGCAGATGAGGAGATCTCTTACCAGAGGAGTCTAACTACAAACCGCTGAAGAGGCGC 1671
 Db 121 GAGCAGATGCGGAAGATCTCTTACCAGAGAGTCTAACTACAAACCGCTGAAGAGGCGC 180
 QY 1672 AAGATGACAAAGTCCATGTTTGTGAAATCAAGACTCTAGGCAATCGGTGCTTTGGGAA 1731
 Db 181 AAGATGACAAAGTCTATGTTTGTCAAGATCAAAACCTGGGATCGGTGCTTTGGGAA 240
 QY 1732 GTGTGCTGCTGTGTAAGCTGGAGACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAG 1791
 Db 241 GTGTGCTGCTGTTGTAAGTGGAGACTCAGCGCTGTACGCCATGAAGAGCCCTTAAGGAAA 300
 QY 1792 AAGATGCTCTGACCGGAAATCAAGTGGCCATGTCAAGCTGAGAGGAGACATCCTGGCT 1851
 Db 301 AAGATGCTCTGACCGGAAATCAAGTGGCCATGTCAAGCTGAGAGGAGACATCCTGGCC 360
 QY 1852 GAAGCAGACAAATGAGTGGGTGTCAAACTCTTACTACTCTCTTCCAAAGACAAAGAGCCTG 1911
 Db 361 GAGCAGACAAATGAGTGGGTGTCAAACTCTTACTACTCTCTTCCAAAGACAAAGAGCCTG 419
 QY 1912 TACTTTGTGATGAGTACATACAGGCGGGGATGATGAGCCTGCTGATCAGGATGAG 1971
 Db 420 TACTTTGTGATGAGTACATACAGGCGGGGATGATGAGCCTGCTGATCAGGATGAG 479
 QY 1972 GTCTTCCCTGAGCAGCTGGCGGCTTCTTACATTCAGAGATGTGACCCCTGGCCATTTGAAGT 2031
 Db 480 GTCTTCCCTGAGCAGCTGGCGGCTTCTTACATTCAGAGATGTGACCCCTGGCCATTTGAAGT 538
 QY 2032 GTCCACAAAGATGGGCTTTATCCACCGGAGATCAAGCCTGACACATACCTCATCGACCTG 2091

```
Db 539 GTCCACAAGATGGCTTATCCACCAGACATCAAGCCTCATAAACATTTTGATAGATCTG 598
QY 2092 GATGGTCATATTAAGCTGACAGATTTGGCCCTGCTGACATGGATTCAGGTGGACTCAAT 2151
Db 599 GATGGTCATATTAAGCTGACAGATTTGGCCCTGCTGACATGGATTCAGGTGGACTCAAT 658
QY 2152 TCCAAGTACTACCAGAAA 2169
Db 659 TTCCAATATTACCAGAAA 676

RESULT 3
US-09-879-536-26/c
; Sequence 26, Application US/09879536
; Patent No. US2002014298A1
; GENERAL INFORMATION:
; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Derti, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/879,536
; PRIOR FILING DATE: 2001-09-21
; PRIOR APPLICATION NUMBER: US 60/088,801
; PRIOR FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO: 26
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-879-536-26

Query Match 10.6%; Score 333; DB 10; Length 638;
Best Local Similarity 71.3%; Pred. No. 1.6e-80;
Matches 454; Conservative 0; Mismatches 180; Indels 3; Gaps 1;

QY 1525 GTCATCAAACTTACCAGCAGAGGTTCAGCGGAGGCTACAGCTGGAGCAGGAAATGGCC 1584
Db 637 GTACTCAAACTTCATCAGCAGCGTCTACATCGTAAAAACAATAGAGAATGAATGATG 578
QY 1585 AAGCTGGGCTGTGTAGGCGCAGGAGGAGAGATGAGGAAGATCTCTACAGAAAGAG 1644
Db 577 CGGGTTGGATTATCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAGAA 518
QY 1645 TCTAATAACAACCGCTGAAGAGGCCAAGATGGACAAGTCCATGTTCTGAAATCAAG 1704
Db 517 TCTAATATACATCCGCTTAAAGAGGCTAAATGGACAAGTCTGTTGTTGAAGATAAG 458
QY 1705 ACTTAGGCTCGGTGCCCTTTGGGAAGTGTGCTGCTGTTGAAGCTGGACACTCACGCT 1764
Db 457 ACATAGGAATAGGAGCATTTGGTGAAGTCTCTAGCAAGAAAGATAGATCTAAGCT 398
QY 1765 CTGTAGGCCATGAAGACTCTCAGGAAGAGGATGCTTGAACCGGAATCAAGTGGCCAT 1824
Db 397 TTGTATGCAACAAAACCTTTCGAAATAAAGATGTTCTTCTCGAAATCAAGTCGCTCAT 338
QY 1825 GTCAAGGCTGAGGGACATCTGCTGAGCAGACAAATGAGTGGTGGTCAACTCTAC 1884
Db 337 GTTAGGCTGAGAGATATCTGCTGAGCTGACAAATGAATGAGTGGTGGTCTCTATAT 278
QY 1885 TACTCCTTCCAGCAAGGACAGCCTGTACTTTGTGTGATGACTACATACCAGCGGGAT 1944
```

```
Db 277 TATTCATCCAAAGATAAGCAAAATTTATACTTTGTATGAATGACTACATTCCTGGGGTGAT 218
QY 1945 ATGATGAGCCTGCTGATCAGGATGGAGGTCTTCCCTGAGACACCTGCCCGCTTCTACATT 2004
Db 217 ATGATGAGCCTATTAATTAAGATGGGCATCTTTCCAGAAAGTCTGGCAGGATTTCTACATA 158
QY 2005 GCAGAGTTGACCTGGCCATTTGAAAGTG--TCCACAAGATGGGCTTTTATCCACCGGGAC 2061
Db 157 GCAGAACTTACCTGTGTCAGTTTGAAAGGTGTTCCCTTAAATGGGTTTATTCATAGAGAT 98
QY 2062 ATCAAGCCTGACAACATACATCTGACCTGGATGTCATATTAAGCTGACAGATTTTGGC 2121
Db 97 ATTAACCTGATAAATATTTTGAATGATCGTATGATGTCATATTAATTAAGTGAATTTGGC 38
QY 2122 CTCTGACCTGGATTCAGGTGGACTCAAAATTCAGT 2158
Db 37 CTCTGACCTGGCTTCAGATGGACACAGGATTCCTAAGT 1

RESULT 4
US-09-938-842A-1861
; Sequence 1861, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Krieps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS, CONTAINING
; FILE REFERENCE: SCRIPT300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 1861
; LENGTH: 1689
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-1861

Query Match 7.3%; Score 231.2; DB 9; Length 1689;
Best Local Similarity 51.8%; Pred. No. 1.2e-52;
Matches 580; Conservative 0; Mismatches 528; Indels 12; Gaps 2;

QY 1504 ATGGAGCAACACCTGGAGAAATGTCATCAAACTTACCAGCAGAAAGTCAAGCGGAGCTA 1563
Db 172 ATCGAATCATTTATAAAATTCAGAAGAGAGTCTCCAGGAAGAAAGAGCGTCGGAGC 231
QY 1564 CAGCTGGAGCAGAAATGCCAAAGCTGGGCTCTGTGAGCGCCGAGCAGGAGATGAGG 1623
Db 232 ATCTTGGAAACAACCTAGCTGCTGATGTACTGTTGAAGACAAGATGATATATTA 291
QY 1624 AAGATCCTCTACCAAGAGAGGTCTAACTACAACCCGGCTCAAGAGGGCCCAAGATGGACAAG 1683
Db 292 AAGAATTTGACAAAAGGAATGAGTATATGCGTCTACAAAGACAGAAATGGGGTT 351
QY 1684 TCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCGCTTTGGGGAAGTGTGCGCTCGCT 1743
Db 352 GATGACTTTGAACTGCTTAGCATCATTTGCCGGGGTGTCTTCGGTGAAGATTTGCT 411
QY 1744 TCTAAGCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGGAGAGAGATGTCTCTG 1803
Db 412 AAGAAAAATCTACTGGAGAGCGCTATATGCAATGAAGAAAGTTAAAGAAATCCGAGATGCTT 471
QY 1804 AACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGGACATCTGCTGGCTGAAGCAGACAAT 1863
```

Db 472 CGACGAGGCGAGTGGACATGTTAAAGCTGAAGAAATGCTTGCGAAGTGGATAGC 531
QY 1864 GAGTGGGGTCAAACTCTACTACTCTCCAGGACAGGACAGCCCTGTGCTACTTTGTGATG 1923
Db 532 CTTTCATTTGCAAGCTTTGCTCTTCCAGATGATGACATTTGTATCTATTATG 591
QY 1924 GACTACATACAGGCGGGGATATGATGAGCCCTGCTGATCAGATGGAGGCTTCCCTGAG 1983
Db 592 GAATACCTCCCTGGAGGTGATGATGACACTGCTGATGCGAAGGATACCTTACGGGAA 651
QY 1984 CACCTGGCCCTCTCTACATTCGACAGTGTGACCTGGCCATTTGAAAGTGTCCACAAGATG 2043
Db 652 GATGAGACTCGTTTATGTTGGCGACACAAATTTCTGCTATTGAGTCTATCCATAAGCAT 711
QY 2044 GGCTTTATCCACCGGACATCAAGCTTCACAACTACTCTACACCTGACCTGGATGTATAT 2103
Db 712 AATTACGTCACAGGATATAAGCCCTGATATTTATGATTAATCTGAAACGCCCATATC 771
QY 2104 AAGCTGACAGATTTGGCCCTCTGACATGGAATTCAGTGGACTCACAAATTTCCAGATTTAAG 831
Db 772 AAGCTTTGAGATTTGGATTTGACAAAGTCTCTGGAAGCAAAATTTCCAGATTTAAG 831
QY 2164 CAGAAAGGAACACATGACAGAGGACATGAGCGCCGCTGACCTGAGAGGCGCA--GAAGCAG 2280
Db 832 CGGAGCTTTGACAGG-----AGTACAAAGCTGACAGAGACATGATAGACTC 882
QY 2224 TCCAACTGCTGCTGGAGACAGGTTTAAAGACCCCTGGAGAGGCGCA--GAAGCAG 2280
Db 883 TCCAAAGGACCTGCTTCTGACCTAGGAGAACTCAGCAGAGACACTTTTACATTTGGCAACA 942
QY 2281 CACGAGAGTGGCTGGACATCTCTTCTGCGGACACCAATATACATCGCTCCGAGGTG 2340
Db 943 AACAGAAGGACCTGCTTCTACAGTAGGAATCCCGATTAATGCTGCTTGGAGCAATCATGTC 1062
QY 2341 CTCTCCGAAAGGATGACGAGCTCTGTGACTGCTGGAGGCTGCTGCTGCTGCTGCTGCT 2400
Db 1003 CTGCTGAAGAAAGGATGGAATGAGTGTGATTTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1062
QY 2401 GAGATGCTGTTGGCAGCGCTTCTTCTGCGGACACCAATATACATCGCTCCGAGGTG 2460
Db 1063 GAGATGCTGAGGTTTCCGCGCTTCTTCCGAGACCTTTCCGAAAGCACTTTGCAACATG 1122
QY 2461 GTGATCACTGGAGAGCAGCTGCTATCTGCTGAGGCTGAGGCTGAGGCTGAGGCTGAGG 2520
Db 1123 ATTTGAACTGGAACCTGCTGAGTTCCTGATGAGCTAGCTCTCCATCGAGGTA 1182
QY 2521 CGAGACCTCATCAGAGCTGCTGCTGCGGCTGACTGCGCGCTGAGGCTGAGGCTGAGG 2580
Db 1183 AAGATCTTATCCGAGACTGCTGCTGCAATGTCGAACAGAGGCTTGGGACCAAGGAGTT 1242
QY 2581 GATGACCTCAAGGACACCCCTTCTTCAACACCATGCACT 2620
Db 1243 CACGAAATCAAGACACACCTTGGTTAGGGGAGTCAAT 1282

RESULT 5
US-09-938-842A-2402
: Sequence 2402, Application US/09938842A
: Patent No. US20020160378A1
: GENERAL INFORMATION:
: APPLICANT: Harper, Jeff
: APPLICANT: Krepis, Joel
: APPLICANT: Wang, Xun
: APPLICANT: Zhu, Tong
: TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
: FILE REFERENCE: SAME, AND METHODS OF USE
: CURRENT APPLICATION NUMBER: US/09938842A
: PRIOR FILING DATE: 2001-08-24
: PRIOR FILING DATE: 2000-08-24
: PRIOR FILING DATE: 2001-01-16

Query Match
Best Local Similarity 7.2%; Score 228; DB 9; Length 1452;
Matches 638; Conservative 0; Mismatches 640; Indels 6; Gaps 2;
PRIOR APPLICATION NUMBER: US 60/300,111
PRIOR FILING DATE: 2001-06-22
NUMBER OF SEQ ID NOS: 5379
SEQ ID NO 2402
LENGTH: 1452
TYPE: DNA
ORGANISM: Arabidopsis thaliana
US-09-938-842A-2402

QY 1433 AGAATAGCAGAGATGAACAGAGAGAGAGTCTCCGATCAAGAGTTACTCCCTTATGCT 1492
Db 23 AGAAGGTACGAGGAGGAGGTTCTGGATCGAGCTTGACCATGGAGAAAGTGGCCGAG 82
QY 1493 TCAATTTCTTCATGGAGCAACACGCTGGAGATGTCTCAAAACCTACACAGAGGTCA 1552
Db 83 CGAAGCAGTACATCGAATCACTACAAAGCTCAGAACAGACATTCAGAGAGGAAG 142
QY 1553 GCCGAGGCTACAGCTGAGCAGGAAATGSCCAAGCTGGGCTCTGTGAGGCGCAGAG 1612
Db 143 AGACACGCTGATCTTAGAAGAAAGTTGGCTCTTCTGAGTGCACAAAGAGAGCAAA 202
QY 1613 AGCAGATGAGAGATCTCTTACGAGAGGAGTCTAACTACACCGCTGAGAGGCGCA 1672
Db 203 TCAACATGATTAAAGACCTGGAGAGAAAGAGACAGAGTTTATGAGGCTTAAAGGA 262
QY 1673 AGATGGACAAAGTCCATTTTGTCAAAATCAAGACTTAGCATTCGCTTTGGGGAAG 1732
Db 263 AGATAGTCTGATGACTTTTGGAGCTTTTGAATATATTTATGCCATGAGAAAGTTAA 382
QY 1733 TGTGCTCTGCTTAAAGCTGGACACTCAAGCTCTGTACCCATGAGAGCTCTCAGAGA 1792
Db 323 TTCCCTTATGCTGGGAGAGAAAGTCTGGAATATTTATGCCATGAGAAAGTTAAAGA 322
QY 1793 AGGATCTCTGACCGGATCAAGTGGCCCTGTCAGAGCTGAGAGGACATCTCTGGCTG 1852
Db 383 CTGAAATGGTCTGAGAGGACAGGTTGAGCATGTGAGAGAGAGAAAGTCTGCTGCTG 442
QY 1853 AAGCAGCAATGAGTGGTGGTCAAACTCTACTCTCTCCAGGACAGACACCTCTGT 1912
Db 443 AGTTTCAAGCCATATATTTGAGAGCTTTACTATTCATTTTCAAGATCCCGAGTATCT 502
QY 1913 ACTTTGTGATGACTACATACAGGCGGGATGATGATGAGCTCTGATCAGGATGAGG 1972
Db 503 ATCTGATTTGAAATATCTCCCGGTGATGATGATGATGATGATGATGATGATGATGAT 562
QY 1973 TCTTCCCTGAGCACCTGCGCTTCTTACATTCGAGAGTTCACCTTGGCCATTTAA 2032
Db 563 CATTACGGGAGATGCTGCCAGATTTTATTTGCTCAAGTGTCTGCTGCTGCTGCTGCT 622
QY 2033 TCCACAAGATGGCTTTTATCCAGCGGACATCAAGCTTGACCAACATCTATCTACCT 2092
Db 623 TACACAGTACAACTATATTTATAGGATATCAAACTGATACCTTCTTTTGGACAAAG 682
QY 2093 ATGTTTATTAAGTACAGATTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2152
Db 683 ATGGCAGATCAAACTCTGAGCTTTGGGCTCTGTAAGGCTCTTGTGTTAGAAATTTAC 742
QY 2153 CCAAGTACTACAGAAAGGAAACCATGATGAGAGGACAGCATGAGGCTTCTGCTGCT 2209
Db 743 CTTCATTTACAGGAGAAATAGGCCACGATGATGAATATGTCAGAACCTATGATGATG 802
QY 2210 TCTGGGACGATTTTCCAACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2269
Db 803 ATAGATCTTTCTGACACTGATAAAGAGAGCTGGGCGAGTCCCGACAGCAAACTTC 862
QY 2270 CGCAGAGCA---GCACAGAGGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2326
Db 863 AGCATTTGGCAGATGAATCGCAGAAACTAGCATTTTCACTGTGGGAACCCGAGCTATA 922

APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/879,536
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 60/088,801
PRIOR FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FASTSEQ for Windows Version 3.0
SEQ ID NO 66
LENGTH: 678
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (1)...(678)
OTHER INFORMATION: n = A,T,C or G
US-09-879-536-66

Query Match
Best Local Similarity 6.8%; Score 214.4; DB 10; Length 678;
Matches 308; Conservative 0; Mismatches 138; Indels 2; Gaps 1;

Qy 1529 TCAAAACCTACCAGCAGAGGCTACGCGGAGGTACAGCTGGAGCAGGAATGCCAAAG 1588
Db 3 TCAAAATCTCATCAGCAGGCTACATCGTAAAAAACAATAGAGAAATGAATGATCGGG 62

Qy 1589 CTGGGCTGTGAGCGCGCAGCAGGAGATGAGGAAGATCCTCTACCAAGAGAGCTA 1648
Db 63 TTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAACTA 122

Qy 1649 ACTCAACCGGCTGAAGAGGCGCCAGATGGACAAGTCCATGTTTGTGAAAATCAAGACT 1708
Db 123 ATTACATCCGCTTAAAGAGGCTAAATGAGCAAGTCTATTTGTGAAGATAAGACAC 182

Qy 1709 TAGGATCGGTGCTTTGGGAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1768
Db 183 TAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGATGATGATTAAGCTTGT 242

Qy 1769 ACGCCATCAAGACTCTCAGGAAGAGGATGCTCTGAAACCGGAATCAAGTGGCCCATGCA 1828
Db 243 ATGCAACAAACACTCTTCAAGAAAGATGCTCTTCTCGAAATCAAGTCTCATGTTA 302

Qy 1829 AGGCTGAGAGGAGACATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1888
Db 303 AGGCTGAGAGAGATATCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 362

Qy 1889 CTTTCCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1948
Db 363 CATTTCCAGATAGAGGC--CATTTTCCTTGTATGGCTACATCTCTTGGGGTATAGA 420

Qy 1949 TGAGCTGCTGATCAGGATGGAGTCTT 1976
Db 421 AGAGCCCATTTAATTAATGGGCATCTT 448

RESULT 8
US-09-764-868-196
Sequence 196, Application US/09764868
Patent No. US2002016871A1
GENERAL INFORMATION:
APPLICANT: Rosen et al.
TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
FILE REFERENCE: PT232
CURRENT APPLICATION NUMBER: US/09/764,868
CURRENT FILING DATE: 2001-01-17
Prior application data removed - refer to PALM or file wrapper
NUMBER OF SEQ ID NOS: 1510
SOFTWARE: PatentIn Ver. 2.0

SEQ ID NO 196
LENGTH: 734
TYPE: DNA
ORGANISM: Homo sapiens
US-09-764-868-196

Query Match
Best Local Similarity 6.1%; Score 192.2; DB 9; Length 734;
Matches 424; Conservative 1; Mismatches 294; Indels 36; Gaps 2;

Qy 1747 AAGCTGGACATCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAGATCTCTCTGAC 1806
Db 12 AAGAAGATACAGGCCATATCTATCAATGAAGATATTGAGAAAGTCTGATGCTTGA 71

Qy 1807 CGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGACATCTGGCTGAAGCAGCAATGAG 1866
Db 72 AAGAGCAGGTGGCCCATATCCGAGCAGAAAGATATTTGGTAAAGCAGCATGGTGC 131

Qy 1867 TGGGTGGTCAAACTCTACTACTCTCTCCAGGACAGGAGCCTGTACTTTGTGTGAC 1926
Db 132 TGGGTGGTGAAGATGTTTACAGTCTTTCAGGATAGAGGAATCTTTATCTAATCATG 191

Qy 1927 TACATACAGGCGGGGATATGATGAGCCTGCTGATCAGGATGAGGCTCTTCCCTGAG 1986
Db 192 TTTCTCCCTGGAGGTGACATGATGACATTTGCAATTAAGAGAAAGACWCTTGTGAC 251

Qy 1987 CTGGCCCGCTTCTACATTCAGAGTTCAGCCCTGGCCATTTGAAAGTGTCCAAAGATGG 2046
Db 252 GAACACAGTCTTACATTCAGAGACTGTTTGGCAATAGATGCGATCCACCAAGTTGG 311

Qy 2047 TTTATCCAGCGGACATCAAGCCTGACACATCTCATCGACCTGGATGCTCATATTAAG 2106
Db 312 TTTATCCATCGGATATTAAAGCCAGACACCTTTTATGGATGCCAAGGCTCATGTAAA 371

Qy 2107 CTGACAGATTTTGGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2166
Db 372 TTTATCTGATTTTGGTATTGATCGGATTAAGAAAGCTCACAGGACTGAATTTTATAG 431

Qy 2167 AAAGGGAAACCATGACAGAGGACAGAGGAGCCGCTGCTGCTGCTGCTGCTGCTGCT 2226
Db 432 AATCTCACACAC-----AACCCACCAAGTGAATTTCTCATTCAGAACATGAATC 481

Qy 2227 AACTGCTGCTGTGAGACAGGTTAAAGACCTTGAGGAGGCGGAGAGCAGCAGCAG 2286
Db 482 AAAGAGAAAGAGCAAACTTG-----GAAGAAAGACAGG 515

Qy 2287 AGGTGCTGCGACATCTCTTGTGGGACACCAAAATFACATCGCTCGGAGGTCTTCT 2346
Db 516 AGCAACTGGCATATCCACAGTTGGGACACAGATTTACATTTGCTCCAGAAATTCATG 575

Qy 2347 CGCAAGGTTACAGCAGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2406
Db 576 CAGACTGGTTACAAACAAATTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 635

Qy 2407 CTGCTGGGAGCGCGCTTTCTTGGCCCGCCACCCACAGAGAGCAGCAGCTGAAGGTGATC 2466
Db 636 CTAATAGGATATCCACCTTCTGCTCTGAAACACTCAAGAGACATACAGAAAGTGTATG 695

Qy 2467 AACTGGGAGCAGCAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 2501
Db 696 AACTGGAAAGAACTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 730

RESULT 9
US-09-771-161A-89
Sequence 89, Application US/09771161A
Patent No. US2002010811A1
GENERAL INFORMATION:
APPLICANT: LEVINE, et al.
TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
FILE REFERENCE: 802620-2005.1
CURRENT APPLICATION NUMBER: US/09/771,161A
CURRENT FILING DATE: 2001-01-26

```

; PRIOR APPLICATION NUMBER: 09/724,676
; PRIOR FILING DATE: 2000-11-28
; PRIOR APPLICATION NUMBER: 136776
; PRIOR FILING DATE: 2000-06-15
; PRIOR APPLICATION NUMBER: 135619
; PRIOR FILING DATE: 2000-04-12
; NUMBER OF SEQ ID NOS: 273
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 89
; LENGTH: 1818
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-771-161A-89

```

Query Match	5.4k	Score	169.4	DB	10	Length	1818
Best Local Similarity	55.8k	Pred.	No.	7.7e-36			
Matches	343	Conservative	0	Mismatches	271	Indels	1
QY	1565	AGCTGGAGCAGGAATGGCCAAAGCTGGGCTCTCTGAGGCCGACAGGACAGATGAGGA	1624				
Db	735	AGTTAGAAAAGGTGATGGAAGAAGAGCCCTAAAGAATGAGGAGAAACGACTCCGCGGAT	794				
QY	1625	AGATCCCTCTACCAAGGAGGTCTAACTACAAACCGGCTGAAGAGGCCAAGATGACAACT	1684				
Db	795	CAGCACATGCTCGGAAGAAACAGAGTTCTTCTCTTTGAAGAGAAACAGACTTGGATTGG	854				
QY	1685	CCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGGAAAGTGTGCCTCGCTT	1744				
Db	855	AAGATTTTGAGTCCCTTAAAGTAATAGCAGAGGAGCATTTGGTGAGTACGGCTTGTTG	914				
QY	1745	GTRAGCTGGACATCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGATGTCTGA	1804				
Db	915	AGAAGAAGAATACGGGACATGTGTATGCAATGAAATACTCCGTAAAGCAGATATGCTTG	974				
QY	1805	ACCGGAATCAAGTGGCCCATCTCAAGCTGAGAGGGACATCCTGGCTGAACGACACATG	1864				
Db	975	AAAAGAGCAGGTGGGCCATCTCGTGGGAGCGTGACATCTAGTGGAGCGACACAGTT	1034				
QY	1865	AGTGGGTGGTCAAACTCTACTCTCTCCAGGACAAGGACAGCCTGTACTTTGTGATGG	1924				
Db	1035	TGTGGGTGTGAAATGTTCTATAGTTTTCAGGATAAGCTAAACCTCTACCTAATCATGG	1094				
QY	1925	ACTACATACAGCGGGGATATGATGACCTGTGTATCAGGATGGAGGTCTTCCCTGAGC	1984				
Db	1095	AGTCTYTGCTGGAGGGGACATGATGACCTGTGTGATGAAAAGACACTCTGACGAAG-	1153				
QY	1985	ACCTGGCCGCTTCTACATTCGACAGTTTGACCTGGCCATTGAAGTGTCCACAAATGG	2044				
Db	1154	AGGAGACTCAGTTTATATACGAGAAACAGATTAGCCCTAGACTCTATTCCAAACCTG	1213				
QY	2045	GCTTTATCCACGGGACATCAAGCCTCACAACTACTCATCGACCTGGATGGTCATATTA	2104				
Db	1214	GATTATCCACAGAGACATCAACCCAGACAACTCTCTTTGGACAGCAAGGGCCATGTA	1273				
QY	2105	AGCTGACAGATTTTGGCCTCTGCACCTGGATTCAGGTGGACTCACAATTTCCAAGTACTACC	2164				
Db	1274	AACTTTCTGACTTTGGTCTTGGACAGGACTGAAAAGACACATAGGCACAGAAATTTATA	1333				
QY	2165	AGAAAGGGAACACCA	2179				
Db	1334	GGAACTGTGAACACCA	1348				

```

RESULT 10
US-09-880-107-3691
; Sequence 3691, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Fockley, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression P

```

```

; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107
; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 3691
; LENGTH: 2549
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 x07767
; NAME/KEY: unsure
; LOCATION: (1)..(2549)
; OTHER INFORMATION: n = a or c or g or t
; US-09-880-107-3691

```

Query Match	5.1%	Score 161.8	DB 10	Length 2549	
Best Local Similarity	59.8%	Pred. No. 1.1e-33			
Matches 271	Conservative 0	Mismatches 182	Indels 0	Gaps 0	
QY	1689	GTTTGTCAAAATCAAGACTCTAGGCATCGGTGCCTTTGGGGAAAGTGTGCCCTCGCTGTAA	1748		
DB	209	GTTTGAACGAATCAAGACCTCGCACGGGCTCTCGGGCGGTGATGCTGGTCAACA	268		
QY	1749	GCTGGACACTCACGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGGATGCTCTGAACCG	1808		
DB	269	CAAGGAGCCGGGAACCACTATGCCATGAAGATCCTCGACAAAGAAGGTGGTGAACCT	328		
QY	1809	GAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCCTGGCTCAAGCAGACAATGAGTG	1868		
DB	329	GAACAGATCGACACACACCCTGAATGAAGACGGCATCTGCAACCTGTCACTTTCGCTT	388		
QY	1869	GGTGGTCAAACTTACTTACTCTTTCAGAGACAAGGACAGCCTGTACTTTGTGATGGACTA	1928		
DB	389	CCTCGTCAAACTCGAGTTCCTTTCAGAGACAACCTAAACTTATACATGGTCAATGGAGTA	448		
QY	1929	CATACAGGCGGGATATGATGACCTGCTCATCAGATGGAGTCTTCCCTGAGCACCT	1988		
DB	449	CGTGCCCGGGGGAGATGTTCTCACACTACGCGCGATCGGAAGTTTCACTGAGCCCCA	508		
QY	1989	GGCCCGCTTCTACATTCGACAGTTGACCTGGCCATTGAAAGTGTCCACAAGATGGGCTT	2048		
DB	509	TGCCCGTTTCTACGCGGCCAGATCGTCTTGACCTTTGAGTATCTGCACCTCGTGATCT	568		
QY	2049	TATCCACCGGCATCAAGCCTGACACATACCTACTCATGCTGGATGGTCAATTAAGCT	2108		
DB	569	CATCTACAGGACCTGAAGCCGGAGATCTGCTCATTTGACCAGCAGGCGCTACATTCAGGT	628		
QY	2109	GACAGATTTGGCCTCTGCACTGGATTCAGGTG	2141		
DB	629	GACAGACTCGGTTTCGCCAAGCGCGTGAAGGG	661		

RESULT 11
US-09-880-107-2340
Sequence 2340, Application US/09880107
Patent No. US20020142981A1
GENERAL INFORMATION:
APPLICANT: Horne, Darci T.
APPLICANT: Vockley, Joseph G.
APPLICANT: Scherf, Uwe
APPLICANT: Gene Logic, Inc.
TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
FILE REFERENCE: 44921-5028-WO
CURRENT APPLICATION NUMBER: US/09/880,107
CURRENT FILING DATE: 2001-06-14
PRIOR APPLICATION NUMBER: US 60/211,379
PRIOR FILING DATE: 2000-06-14
PRIOR APPLICATION NUMBER: US 60/237,054

Best Local Similarity 56.5%; Pred. No. 2.3e-30; Mismatches 227; Indels 3; Gaps 1; Matches 299; Conservative

QY 1602 GCCGAGCAGGAGGATGAGGAAGATCTCTTACCAGAGGAGTCTTAACCTACAACCGGCT 1661
DB 899 GCCCAGGACAAAGTACGTGGCGGACTTCTTGCAGTGGCGGAGCCCATCGTGTGGAGGCT 958
QY 1662 GAAGAGGCCCAAGATGGACAAAGTCCATCTTTGTAAGTCTTGTGAAATCAAGACTCTAGGCATCGGTGC 1721
DB 959 TAAGGAGGTCCGACTGAGAGGGGAGGAGTCTGAGATTTCTGAGGTGATCGAGCGGGGC 1018
QY 1722 CTTTGGGAAAGTGCCTCGCTTGTAAAGTCTGAGCAGTCTGACGCTCTGTACGCCATGAAGAC 1781
DB 1019 GTTCAGCGAGGTAGCGGTAGTGAAGATGAAGCAGACGCGGAGGTGTATGCCATGAAGAT 1078
QY 1782 TCTCAGGAAGAAGATGTCCTGAACCGGAATCAAGTGGCCCATCTCAAGGCTGAGAGGGA 1841
DB 1079 CATGAACAAGTGGGACATGCTGAAGAGGGGCGAGGTGCTGCTTCCGTGAGGAGAGGGA 1138
QY 1842 CATCTGGCTGAAGCAGACAAATGAGTGGTGGTCAAACTTACTACTCTTCTTCCAGGACAA 1901
DB 1139 CGTGTGGTGAATGGGACCGCGGTGGATCAGCAGCTGCACCTTCGCTTCCAGGATGA 1198
QY 1902 GGACAGCCTGTACTTTGTGATGAGTACATACCAGCGGGGATGATGATGAGCCTGCTGAT 1961
DB 1199 GAACCTACCTGTACCTGGTCAATGAGTATTTACGTGGCGGGGACCTGCTGACACTGCTGAG 1258
QY 1962 CAGGAT---GGAGGTCTTCCCTGAGCACCCTGGCGGCTTCTACATTTGAGAGTTGACCT 2018
DB 1259 CAGTTTGGGAGCGGATTCGCGCGGAGATGCGCGCTTCTACCTGGCGGAGATTGTGAT 1318
QY 2019 GGCATTAAGAGTGTCCCAAGATGGGCTTTATCCAGCGGACATCAAGCCTGACCAACAT 2078
DB 1319 GGCATAGACTCGGTGACCGGCTTGGCTACGTGACAGGGGACATCAACCCGACCAACAT 1378
QY 2079 ACTCATCGACCTGGATGGTCAATTAAGCTGACAGATTTTGGCCTCTGC 2127
DB 1379 CCGTGGACCGCTGTGGCCACATCCGCTGGCGGACTTCGCTCTTGC 1427

RESULT 13
US-09-938-842A-633
; Sequence 633, Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIPI300-3
; CURRENT APPLICATION NUMBER: US/09/938,842A
; PRIOR FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 633
; LENGTH: 1398
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-633

Query Match 4.5%; Score 142; DB 9; Length 1398;
Best Local Similarity 57.4%; Pred. No. 1.8e-28;
Matches 256; Conservative 0; Mismatches 190; Indels 0; Gaps 0;

QY 1690 TTTGTGAAATCAAGACTCTAGGCATCGTCTTTGGGAGGAGTGTGCTCGCTTGTGAAG 1749

Best Local Similarity 4.9%; Score 155.6; DB 10; Length 1635;
Best Local Similarity 59.5%; Pred. No. 4e-32; Mismatches 179; Indels 0; Gaps 0; Matches 263; Conservative

QY 1700 TCAAGACTCTAGGATCGGTGCTTTGGGAAAGTGGCTCGCTTGTGAAGCTGGACACTC 1759
DB 192 TCAGGACGCTGGGATGGGCTCTTCGGGGGGGTGATGCTGGTGAGGCACCGAGACCG 251
QY 1760 AGGCTCTCTACGCCATGAAGACTCTCAGGAAGAAGGATGCTCTGAACCGGAATCAAGTGG 1819
DB 252 GCGGCCACTACGCCATGAAGTCTCAGAGAGAGAGGTGGTGAAGATGAAGCAGGTGG 311
QY 1820 CCATGTCTCAAGCTGAGAGGACATCTCTGAGTGAAGCAGACAAATGAGTGGTGTCAAAAC 1879
DB 312 AGCACATCTGAACGAGAGGCGATCTCTGAGGCGGATGCGACTTTCCGTTCTCTGTCAGC 371
QY 1880 TCTACTACTCTCCAGCAGAGGACGCTGTACTTTGTGATGAGTACTACATACAGGCG 1939
DB 372 TCCAGTCTCTTAAAGACAACTCTTACCTGTACCTGTGAGTACGTGCGGGTG 431
QY 1940 GGGATATGATGAGCTGCTGTATGAGTGGAGGTCTTCCCTGAGCACCCTGCCCGCTTCT 1999
DB 432 GGGAGATGTTCTCCCGCTTACAGCGGCTCGGAAGTGTAGCGAGCGCCATCGCTTCT 491
QY 2000 ACATTCAGAGTGTACCTCGGCGCATGAAGTGTCCCAAGATGGCGTTTATCCACCGGG 2059
DB 492 ATGCCCGCCAGGTGCTTGGCGGCTCCAGTACCTACATCTGCTGACCTATCCAGCGCG 551
QY 2060 ACATCAAGCTGACACATCTACTGAGCTGGATGGTCAATTAAGCTGACAGATTTTG 2119
DB 552 ACCTGAAGCCGAGATCTCTCATCGACGAGGCTTACCTGAGGTGAGGACTTCG 611
QY 2120 GCCTCTGCACTGATTCAGGTG 2141
DB 612 GTTTCGCCAAGCGCGTGAAGGG 633

RESULT 12
US-09-971-845-1
; Sequence 1, Application US/09971845
; Patent No. US20020132247A1
; GENERAL INFORMATION:
; APPLICANT: DELANEY, ALLEN
; APPLICANT: YOGANATHAN, THILLAINATHAN
; TITLE OF INVENTION: DYSTROPHIA MYOTONICA PROTEIN KINASE
; FILE OF INVENTION: (DM-PK) AND ITS USES
; FILE REFERENCE: KINEO28CON
; CURRENT APPLICATION NUMBER: US/09/971,845
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/21479
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/238,558
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3407
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-971-845-1

Query Match 4.7%; Score 149.8; DB 10; Length 3407;

Db 400 TTTGAGTTATGAAGCTTCTGGGAAAGGTGGCTTTGGGAAAGCTTACCAAGGTGAGGAA 459
Qy 1750 CTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTCTGAACCGG 1809
Db 460 AAGGAGACCTCTGAGATATACGCTATGAAGTCAATCAGGAAGATCAGATATATGGAGAG 519
Qy 1810 AATCAAGTGGCCATGCTCAAGGCTGAGAGGGACATCCTGGTGAAGCAGCAATGAGTGG 1869
Db 520 AACCATGCTGAATACATGAAGCTGAGCGGATATCTTTACAAAAATTGATCACCATTTC 579
Qy 1870 GTGGTCAAACTCTACTACTCTTCCAGGACAAGGACGCTGTACTTTTGTGATGGACTAC 1929
Db 580 ATGTTCAACTTAAATACTCTTTTCAGACCAAGTACAGGCTGTATCTTGTGCTGACTTT 639
Qy 1930 ATACCAAGCGGGGATGATGAGCTCTGCTGATCAGGATGAGGCTCTTCCCTGAGCACCTG 1989
Db 640 ATAAAGCGAGGTCACTCTTTTCTTCCAACTCTATCACCAGGCTTTTCAGAGAGGACTTG 699
Qy 1990 GCGCGCTCTTACATGTCAGAGTTGACCCCTGGCCATTTGAAAGTGTCCACAAAGTGGCTTT 2049
Db 700 GCTGAGTGTACACTGACAGAAATCGTCTCTGAGTTTCCCATCTCCATGAGAAAGGCATA 759
Qy 2050 ATCCACCGGAGCATCAAGCTTGACACATACATCAGCTGATGAGGCTCTTCCCTGAGCACCTG 2109
Db 760 ATGCACAGAGATCTGAAACCCGAAACATACATGATGATACAGATGCCACCGTATGCTTA 819
Qy 2110 ACAGATTTTGGCTCTGCACTGGATT 2135
Db 820 ACAGATTTTGGTTAGCAAGAAATT 845

RESULT 14

US-09-938-842A-2503

; Sequence 2503, Application US/09938842A

; Patent No. US20020160378A1

; GENERAL INFORMATION:

; APPLICANT: Harper, Jeff

; APPLICANT: Kreps, Joel

; APPLICANT: Wang, Xun

; APPLICANT: Zhu, Tong

; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING

; FILE REFERENCE: SAME, AND METHODS OF USE

; CURRENT APPLICATION NUMBER: US/09/938,842A

; PRIOR FILING DATE: 2001-08-24

; PRIOR FILING DATE: 2000-08-24, 666

; PRIOR FILING DATE: 2000-08-24, 647

; PRIOR FILING DATE: 2001-01-16

; PRIOR FILING DATE: 2001-06-22

; NUMBER OF SEQ ID NOS: 5379

; SEQ ID NO 2503

; LENGTH: 1416

; TYPE: DNA

; ORGANISM: Arabidopsis thaliana

US-09-938-842A-2503

Query Match 4.4%; Score 138.8; DB 9; Length 1416;
Best Local Similarity 57.0%; Pred. No. 1.3e-27;
Matches 254; Conservative 0; Mismatches 192; Indels 0; Gaps 0;

Qy 1690 TTTGTGAAATCAAGACTCTAGGCATCGCTTTGGGAAAGTGTGCTCTGCTGTAAG 1749
Db 418 TTTGAGTTATGAAGTTGTGGGCAAGGTGCTTTGAAAGTGTACCAAGTGGAGGAA 477
Qy 1750 CTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTCTGAACCGG 1809
Db 478 AAAGACACCTCTGAGATATACCGATGAAGTCAATGAGAAAGATAAATTTCTTGAGAG 537
Qy 1810 AATCAAGTGGCCATGCTCAAGGCTGAGAGGGACATCTGCTGGCTGAAGCAGCAATGAGTGG 1869
Db 538 AATCATGCTGAATACATGAAGCCGAGCGGATATTTCAACCAAAATCATCATCTTTC 597

Qy 1870 GTGGTCAAACTCTACTACTCTTCCAGGACAAGGACAGCTGTACTTTTGTGATGGACTAC 1929
Db 598 ATGTGCAACTTAAATACTCTTTTCAGACCAATACAGATTGTATCTTCTTGTGACTTT 657
Qy 1930 ATACCAAGCGGGGATATGATGAGCCCTGTGATCAGGATGAGGCTCTTCCCTGAGCACCTG 1989
Db 658 ATAAAGCGGAGGTCACTCTTTTCTTCCAGCTCTATCACCAGGCTTTTCAGGAGGACTTG 717
Qy 1990 GCGCGCTCTACATTCGACAGTTGACCCCTGGCCATTGAAAGTGTCCACAAGATGGCTTT 2049
Db 718 GCTGCTGTACACTGACAGAAATCGTCTCTGAGTTTCCCATCTCCATGAGAAAGCATA 777
Qy 2050 ATCCACCGGAGCATCAAGCTTGACACATACATCAGCTGATGAGTGTGATTAAGCTG 2109
Db 778 ATGCATAGAGATCTTAAACCTGAAACATACATCAGGATGAGTGTGATGATTA 837
Qy 2110 ACAGATTTTGGCTCTGCACTGGATT 2135
Db 838 ACTGATTTTGGTTAGCAAGAAATT 863

RESULT 15

US-09-764-868-58

; Sequence 58, Application US/09764868

; Patent No. US20020168711A1

; GENERAL INFORMATION:

; APPLICANT: Rosen et al.

; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies

; FILE REFERENCE: PT232

; CURRENT APPLICATION NUMBER: US/09/764,868

; CURRENT FILING DATE: 2001-01-17

; Prior application data removed - refer to PALM or file wrapper

; NUMBER OF SEQ ID NOS: 1510

; SOFTWARE: Patentin Ver. 2.0

; SEQ ID NO 58

; LENGTH: 1735

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-764-868-58

Query Match 4.3%; Score 135.6; DB 9; Length 1735;
Best Local Similarity 58.3%; Pred. No. 1.1e-26;
Matches 253; Conservative 2; Mismatches 176; Indels 3; Gaps 1;

Qy 1690 TTTGTGAAATCAAGACTCTAGGCATCGCTTTGGGAAAGTGTGCTCTGCTGTAAG 1749
Db 297 TTTGAAATTAATTAAGTAATTTGGAAGAGGTGCTTTTGGTGAAGTGTCTTGTCAAAATG 356
Qy 1750 CTGGACACTCAGCTCTGTACGCCATGAAGACTCTCAGGAAGAGGATGCTCTGAACCGG 1809
Db 357 AAGATACTGAAGAAATTTATGCAATGAATAATCTCAACAAGTGGAGATGCTGAAGA 416
Qy 1810 AATCAAGTGGCCATGTCAAGGCTGAGAGGGACATCTCTGGCTGAAGCAGACAATGAGTGG 1869
Db 417 GCAGAGACCGCTGCTTCCGAGAGGAGCGCATGTGCTGTCACACGCGACTGCCAGTGG 476
Qy 1870 GTGGTCAAACTCTACTACTCTTCCAGGACAAGGACAGCTGTACTTTTGTGATGGACTAC 1929
Db 477 ATCACCAGGCTCTACATCGCTTTTCAGGAGGAGAACCCACTGTACTTGTATGATGATTAC 536
Qy 1930 ATACCAAGCGGGGATATGATGAGCTGTGATCAGGATGAGGCTCTTCCCTGAGCAAC 1986
Db 537 ATGTGGTGGTGTGATTTACTGACCTCTGCTCASCATAATTTGAAGACAAGCTTCGGAAGAT 596
Qy 1987 CTGGCCCGCTTCTACATTTGACAGATTGACCCCTGGCCATTGAAAGTGTCCCAAGATGGGC 2046
Db 597 ATGCGGAGTTTCTACATTTGGTGAATTTGGTGTGCTGGCCATTGACTCCATCCATCAGCTCAT 656
Qy 2047 TTTATCCAGCGGACATCAAGCTTGACAACTACTCATCTGACCTGGATGGTGTATTAAG 2106
Db 657 TACGKGACAGRACATTAACCTGACATGTCTCTTTGGAGCGTGAATGCTCATATCCSC 716

Qy 2107 CTGACAGATTGG 2120
||| | || ||||
Db 717 CTGGCTGACTGG 730

Search completed: January 16, 2003, 22:04:38
Job time : 144.552 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:55:17 ; Search time 2727.25 Seconds
(without alignments)
18735.612 Million cell updates/sec

Title: US-09-763-334-5
Perfect score: 3155
Sequence: 1 atgagagccacccgaagtt.....aagagcattatttggggg 3155

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0
Searched: 16154066 seqs, 8097743376 residues
Total number of hits satisfying chosen parameters: 32308132

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
EST:
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_hic:*
9: gb_est1:*
10: gb_est2:*
11: gb_hic:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_Other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	ID	Description
1	927.6	29.4	1547	11 BC024819	BC024819 Mus muscu
2	782.2	24.8	926	13 BI412714	BI412714 602987439
3	781	24.8	924	14 BQ930443	BQ930443 AGENCOURT
4	772	24.5	875	13 BI664597	BI664597 603290249
5	725	23.0	737	14 BM950668	BM950668 UI-M-EH09
6	659.8	20.9	687	10 BB354465	BB354465 BB354465

7	637.4	20.2	930	14 BQ898648	BQ898648 AGENCOURT
8	622	19.7	689	13 BI331257	BI331257 602982912
9	578.8	18.3	824	13 BI525737	BI525737 602926563
10	562.4	17.8	575	12 BE986745	BE986745 UI-M-CG0P
11	553.8	17.6	569	12 BF461270	BF461270 UI-M-CG0P
12	540.6	17.1	549	10 AW494954	AW494954 UI-M-BH3-
13	537.4	17.0	539	12 BF465780	BF465780 UI-M-CG0P
14	533.8	16.9	537	10 AW456717	AW456717 UI-M-BH3-
15	519.4	16.5	863	10 BE620135	BE620135 601483053
16	508.6	16.1	515	13 BI133909	BI133909 UI-M-BH3-
17	503.6	16.0	590	13 BM052863	BM052863 1667C07.Y
18	497.6	15.8	631	14 BM801311	BM801311 AGENCOURT
19	494.8	15.7	736	12 BG538495	BG538495 602567216
20	494.2	15.7	579	13 BI964175	BI964175 1665909.Y
21	479.4	15.2	596	12 BF137690	BF137690 601782626
22	476.4	15.1	478	12 BF460475	BF460475 UI-M-CG0P
23	471.4	14.9	473	12 BE987358	BE987358 UI-M-CG0P
24	458.8	14.5	462	12 BF463863	BF463863 UI-M-CG0P
25	458.8	14.5	462	12 BF465982	BF465982 UI-M-CG0P
26	443.6	14.1	462	12 BF465982	BF465982 UI-M-CG0P
27	442	14.0	450	9 AI430896	AI430896 mh70d12.x
28	438.6	13.9	445	10 AW493915	AW493915 UI-M-BH3-
29	429	13.6	686	10 BB614487	BB614487 BB614487
30	420.2	13.3	447	9 AA023402	AA023402 mh71d12.r
31	403.4	12.8	416	14 W81967	W81967 me93h03.r1
32	395.4	12.5	397	12 BF464894	BF464894 UI-M-CG0P
33	394.2	12.5	528	14 BM706472	BM706472 UI-E-CR1-
34	382.8	12.1	460	10 AW227865	AW227865 up15c06.Y
35	362	11.5	648	13 BI155504	BI155504 602903902
36	352	11.3	468	12 BF774005	BF774005 283544 MA
37	346.2	11.0	445	10 BB780786	BB780786 BB780786
38	343.8	10.9	355	10 AW909686	AW909686 uc72g09.Y
39	343	10.9	602	9 AL601079	AL601079 DKF2P313K
40	341.2	10.8	962	17 CNS0396F	AL233520 Tetraodon
41	338.2	10.7	343	12 BE853641	BE853641 ux20a08.Y
42	338	10.7	442	10 AW653741	AW653741 102578 MA
43	335	10.6	1053	14 U69566	U69566 U69566 Soar
44	333.4	10.6	488	10 BE289622	BE289622 601092179
45	332.6	10.5	419	10 AV655751	AV655751 AV655751

ALIGNMENTS

RESULT 1
BC024819
LOCUS BC024819 1547 bp mRNA linear HTC 07-AUG-2002
DEFINITION Mus musculus, Similar to large tumor suppressor 2, clone
IMAGE:5360539, mRNA.
ACCESSION BC024819
VERSION BC024819.2 GI:20380582
KEYWORDS HTC.
SOURCE house mouse.
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE
1 (bases 1 to 1547)
Strausberg,R.
Direct Submission
Submitted (01-MAR-2002) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer
Institute, 31 Center Drive, Room 11A03, Bethesda, MD 20892-2590,
USA
NIH-MGC Project URL: <http://mgc.nci.nih.gov>
On May 1, 2002 this sequence version replaced gi:19354132.
COMMENT
Contact: MGC help desk
Email: cgapbs-r@mail.nih.gov
Tissue Procurement: The Cepko Laboratory
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Baylor College of Medicine Human Genome
Sequencing Center
Center code: BCM-HGSC

Web site: <http://www.hgsc.bcm.tmc.edu/cdna/>
 Contact: ang@bcm.tmc.edu
 Gunaratne, P.H., Garcia, A.M., Lu, X., Hulyk, S.W., Hale, S.M.,
 Yoon, V.S., Kowis, C.R., Lawrence, S., Martin, R.G., Muzny, D.M.,
 Richards, S., Gibbs, R.A.

Clone distribution: MGC clone distribution information can be found
 through the I.M.A.G.E. Consortium/LLNL at: <http://image.llnl.gov>
 Series: IRAK Plate: 54 Row: m Column: 22
 This clone was selected for full length sequencing because it
 passed the following selection criteria: matched mRNA gi: 7657294
 This clone has the following problem: no polyA-tail.

FEATURES

source

Location/Qualifiers
 1..1547
 /organism="Mus musculus"
 /db_xref="taxon:10090"
 /clone="IMAGE:5360539"
 /tissue_type="Eye, retina, mouse strain C57BL/6"
 /clone_lib="NIH_MGC_94"
 /lab_host="DH10B"
 /note="Vector: pCMV-SF0RT6"
 350 a 468 c 446 g 283 t

BASE COUNT
 ORIGIN

Query Match 29.4%; Score 927.6; DB 11; Length 1547;
 Best Local Similarity 99.6%; Pred. No. 1.8e-223;
 Matches 930; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 ATGAGAGCCACCCGAGTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATCC 60
 DB 614 ATGAGAGCCACCCGAGTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATCC 673
 QY 61 CTCCTGCTTTGCCAAGAGTCAGGCACTTCGGCAGCTGCAGAGGTGAACCGGCAGATG 120
 DB 674 CTCCTGCTTTGCCAAGAGTCAGGCACTTCGGCAGCTGCAGAGGTGAACCGGCAGATG 733
 QY 121 CTTACAGGAGTTGGTAATGCGGCATGTGCACAGAGATGCTGCAGAGCCTCACGCAG 180
 DB 734 CTTACAGGAGTTGGTAATGCGGCATGTGCACAGAGATGCTGCAGAGCCTCACGCAG 793
 QY 181 ACGGCAGTAGGAGTATCGAAGCTGCCTTGGAGTACATCAGTAAGTGGGTACCTCGAC 240
 DB 794 ACGGCAGTAGGAGTATCGAAGCTGCCTTGGAGTACATCAGTAAGTGGGTACCTCGAC 853
 QY 241 CCCAGGATGACAGATGTGCGAGTATCATCAAGCAGACCTCCCGAGAAAGGCGCTGGCG 300
 DB 854 CCCAGGATGACAGATGTGCGAGTATCATCAAGCAGACCTCCCGAGAAAGGCGCTGGCG 913
 QY 301 TCCACCCCGGTGACTCGCGGCCCGCAGTTTCGAGGGCACAGGGAAGCACTCCCATCTAC 360
 DB 914 CCCACCCCGGTGACTCGCGGCCCGCAGTTTCGAGGGCACAGGGAAGCACTCCCATCTAC 973
 QY 361 CACCAGCTGGGTGTGCAAACTACGAGGGCCCGCCGCACTGGAGAGATGCCCGGCA 420
 DB 974 CACCAGCTGGGTGTGCAAACTACGAGGGCCCGCCGCACTGGAGAGATGCCCGGCA 1033
 QY 421 TATTAGACTTCTCTTCCTCGGAGCGGAGCGGACCCACCGTCCAGAGGCTCACCA 480
 DB 1034 TATTAGACTTCTCTTCCTCGGAGCGGAGCGGACCCACCGTCCAGAGGCTCACCA 1093
 QY 481 CATCTCCCAAGAGGTACAGCAGCAGTAGAGCCAAAGTGGCACTTTCGGGSCACACAC 540
 DB 1094 CATCTCCCAAGAGGTACAGCAGCAGTAGAGCCAAAGTGGCACTTTCGGGSCACACAC 1153
 QY 541 TATGGTGTGGTGTATCTATCTATCGGAGCAGTCTGGGTATGGGTGCAGCGCAGTTCTCC 600
 DB 1154 TATGGTGTGGTGTATCTATCTATCGGAGCAGCCTGGGTATGGGTGCAGCGCAGTTCTCC 1213
 QY 601 TTCAGAAACAGAGCCGACAGATGCTTATCCAGATGCCAAGCCGAGGTTGGCCCT 660
 DB 1214 TTCAGAAACAGAGCCGACAGATGCTTATCCAGATGCCAAGCCGAGGTTGGCCCT 1273
 QY 661 CCGCGCAGCTACCTTTCTCTCCGCCATGCTGGGCTGTATACACTGCCTCGCACCAAGCGC 720

DB 1274 CCCGCCACCCCTACCTTTCTGCCCCATCTGGGTGTACACTGCTGCACCAAGCG 1333
 QY 721 CGGGCTACCCACCTGGGGCCACCCATTACATGTGTGGGCACCCGGGTCACCGCTTT 780
 DB 1334 CGGGCTACCCACCTGGGGCCACCCATTACATGTGTGGGCACCCGGGTCACCGCTTT 1393
 QY 781 ACTGGCGAAGAGCTTGCACAGGCTGTGCTGCACCGCTGCAGGAACAGCCTCAATGCTGAC 840
 DB 1394 ACTGGCGAAGAGCTTGCACAGGCTGTGCTGCACCGCTGCAGGAACAGCCTCAATGCTGAC 1453
 QY 841 TTGTACGAGCTGGCTCCACGCTGCCTGCTGCACCTCCACTCCACTGGCAGCGGACTCG 900
 DB 1454 TTGTACGAGCTGGCTCCACGCTGCCTGCTGCACCTCCACTCCACTGGCAGCGGACTCG 1513
 QY 901 CTCGAGAAGCAGGCTCTAGAAGCCTCGCGCGCGC 934
 DB 1514 CTCGAGAAGCAGGCTCTAGAAGCCTCGCGCGCGC 1547

BI412714 926 bp mRNA linear EST 14-AUG-2001
 602987439F1 NCI_CGAP_Lu33 Mus musculus cDNA clone IMAGE:5143496 5',
 mRNA sequence.
 BI412714
 BI412714.1 GI:15173637
 EST.
 house mouse.
 Mus musculus
 Mammalia; Euthera; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Metazoa; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 NIH-MGC <http://mgi.nci.nih.gov/>
 National Institutes of Health, Mammalian Gene Collection (MGC)
 Unpublished (1999)
 Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Gilbert Smith, Ph.D.
 cDNA Library Preparation: M. Bento Soares, Ph.D., M. Fatima
 Bonaldo, Ph.D.
 cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Clone Sequencing by: Incyte Genomics, Inc.
 DNA distribution: NCI-CGAP clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Plate: LLAM1353 row: n column: 09
 High quality sequence start: 4
 High quality sequence stop: 858.
 Location/Qualifiers
 1..926
 /organism="Mus musculus"
 /strain="CZECH II"
 /db_xref="taxon:10090"
 /clone="IMAGE:5143496"
 /clone_lib="NCI_CGAP_Lu33"
 /tissue_type="pooled lung tumors"
 /lab_host="DH10B (phage-resistant)"
 /note="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
 modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
 strand cDNA was prepared from mRNA obtained from pooled
 lung tumors with a Not I - oligo(dT) primer [5',
 TGTACCAATCTGAAGTGGAGGCGGCGCTCTGTCTTTTCTTTT 3']
 Double-stranded cDNA was ligated to Eco RI adaptors
 (Pharmacia), digested with Not I and cloned into the Not
 I and Eco RI sites of the modified pT73 vector. Library
 went through one round of normalization, and was
 constructed by Bento Soares and M. Fatima Bonaldo."

FEATURES

source

BASE COUNT
 ORIGIN

Query Match 24.8%; Score 782.2; DB 13; Length 926;
 Best Local Similarity 97.5%; Pred. No. 9e-187;

Matches 890; Conservative 0; Mismatches 13; Indels 10; Gaps 9;			
QY	9	CACCCGGAAGTTGGACCTTATCAAAAGCTCTCAGGGAATCCGATATTCCTCTCTGCC	68
Db	20	CACCCGGAAGTTGGACCTTATCAAAAGCTCTCA--GGAAATCCGATATTCCTCTCTGCC	78
QY	69	TTTGGCAACAGCTCAGGACCTTGGCAGCTGCAGAGTGAACCGGCAGATGCTTCAGGA	128
Db	79	TTTGGCAACAGCTCAGGACCTTGGCAGCTGCAGAGTGAACCGGCAGATGCTTCAGGA	137
QY	129	CTTGTGAATGGGCGATGTACACAGAGATGCTGCAGAGGCTTACCCAGACGGGCGAG	188
Db	138	CTTGTGAATGGGCGATGTACACAGAGATGCTGCAGAGGCTTACCCAGACGGGCGAG	197
QY	189	TAGGAGTATCGAAGCTGCTTGGAGTACATCAGTAGATGGCTACCTGGACCCAGGAA	248
Db	198	TAGGAGTATCGAAGCTGCTTGGAGTACATCAGTAGATGGCTACCTGGACCCAGGAA	257
QY	249	TGAGCAGATTGTGCGAGTCAATCAAGCAGACCTTCCCGAGGAAGGGCTTGGCGTCCACCC	308
Db	258	TGAGCAGATTGTGCGAGTCAATCAAGCAGACCTTCCCGAGGAAGGGCTTGGCGTCCACCC	317
QY	309	GCTGACTCGCGGCGCCAGTTTTCAGGGGCACAGGGGAAGCACTCCCATCTACACAGCT	368
Db	318	GCTGACTCGCGGCGCCAGTTTTCAGGGGCACAGGGGAAGCACTCCCATCTACACAGCT	377
QY	369	GGTGTGTCAAACTACGAGGGCCCGCCGCTGAGGAGATGCCGCGCAATATTTAGA	428
Db	378	GGTGTGTCAAACTACGAGGGCCCGCCGCTGAGGAGATGCCGCGCAATATTTAGA	437
QY	429	CTTCTCTTCTTCCCTGGAGCGCGGCGCCACCCAGCTGCTCCAGGCTCACCAGCTCTCC	488
Db	438	CTTCTCTTCTTCCCTGGAGCGCGGCGCCACCCAGCTGCTCCAGGCTCACCAGCTCTCC	497
QY	489	CAAGGGTACAGCAGCAGTATGAGCCAGTGGGCACCTTTCGGGGCACACATATGGTGG	548
Db	498	CAAGGGTACAGCAGCAGTATGAGCCAGTGGGCACCTTTCGGGGCACACATATGGTGG	557
QY	549	TGCTATCTACTATCGGACAGCTCTGGGTATGGGTGACGCGCAGTTCCTCTTCCAGAA	608
Db	558	TGCTATCTACTATCGGACAGCTCTGGGTATGGGTGACGCGCAGTTCCTCTTCCAGAA	617
QY	609	CAAGAGCCACAGATGCTTATCCAGCATGCGCCAGGCGCCAGGCTGCTCCGCGCAG	668
Db	618	CAAGAGCCACAGATGCTTATCCAGCATGCGCCAGGCGCCAGGCTGCTCCGCGCAG	677
QY	669	CCTCAGCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	727
Db	678	CCTCAGCTTCTTCCCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	737
QY	728	CCCCAGCTGGGCGCCACCATTTACATGTGTGGGACCGCGGGTCCAGCTTACTGGGG	787
Db	738	CCCCAGCTGGGCGCCACCATTTACATGTGTGGGACCGCGGGTCCAGCTTACTGGGG	795
QY	788	AAAGCTCTGCAGAGCTGTGCTGGACCTGTCAGGAAACAGCTCAATGCTGACTGTAGG	847
Db	796	AAAGCTCTGCAGAGCTGTGCTGGCA--CGTCCAGGAACAG--CTCAATGCTGACTGTAGG	853
QY	848	-AGCTGGGCTCCAC--GGTCCCTGGTC--TGCAGCTCCACTGCGACGCGCGGCTGCTGC	904
Db	854	AAAGCTGGGCTCCACGGGTCCTGCTGGTCTTGGAGTCCCATGTCGACGCGCGGAACTGCTGC	913
QY	905	AGAAGCAGGGTCT 917	
Db	914	AGAAGCAGGGTCT 926	

RESULT 3
BO930443
LOCUS
DEFINITION BO930443 924 bp mRNA linear EST 20-AUG-2002
AGENCOURT_8931059 NIH_MGC_94 Mus musculus cDNA clone IMAGE:6468340
5', mRNA sequence.
ACCESSION BO930443

VERSION	BO930443.1	GI:22345474
KEYWORDS	EST.	
SOURCE	house mouse.	
ORGANISM	Mus musculus	
REFERENCE	Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus. 1 (bases 1 to 924)	
AUTHORS	NIH-MGC http://mgi.nci.nih.gov/	
TITLE	National Institutes of Health, Mammalian Gene Collection (MGC)	
JOURNAL	Unpublished (1999)	
COMMENT	Contact: Robert Strausberg, Ph.D. Email: cgabbs-remail.nih.gov Tissue Procurement: The Cepko Laboratory cDNA Library Preparation: Life Technologies, Inc. DNA Sequencing by: Agencourt Bioscience Corporation Clone distribution: MGC clone distribution information can be found through the I.M.A.G.E. Consortium/LLNL at: http://image.llnl.gov plate: LLAM13995 row: 1 column: 05 High quality sequence stop: 554.	
FEATURES	Location/Qualifiers	
source	1..924	
	/organism="Mus musculus"	
	/db_xref="taxon:10090"	
	/clone="IMAGE:6468340"	
	/clone_lib="NIH_MGC_94"	
	/tissue_type="retina"	
	/lab_host="DH10B (phage-resistant)"	
	/note="Organ: eye; Vector: pCMV-SPORT6; Site: 1: NotI; Site 2: SalI; Cloned unidirectionally; oligo-dT primed. Average insert size 3.3 kb. Library enriched for full-length clones and constructed by Life Technologies. Note: this is a NIH_MGC Library."	
BASE COUNT	230 a	257 c 241 g 196 t
ORIGIN		
Query Match	24.8%	Score 781; DB 14; Length 924;
Best Local Similarity	99.3%	Pred No. 1.8e-186;
Matches 795; Conservative 0; Mismatches 5; Indels 1; Gaps 1;		
QY	2354	GGTACACGAGCTGTGACTGGTGGAGGCTGGTGTGATCTCTTTGAGATGCTGGTGG 2413
Db	7	GGGACACGAGCTGTGACTGGTGGAGGCTGGTGTGATCTCTTTGAGATGCTGGTGG 66
QY	2414	GGCAGCGCTTCTTGGCCCCCCCCCCCCACAGACGAGCTGAGGTGATCACTGGG 2473
Db	67	GGCAGCGCTTCTTGGCCCCCCCCCCCCACAGACGAGCTGAGGTGATCACTGGG 126
QY	2474	AGAGCAGCTGCATATCCCTACGAGGTGAGGCTCAGCGCTCAGGCGCCGAGACCTCATCA 2533
Db	127	AGAGCAGCTGCATATCCCTACGAGGTGAGGCTCAGCGCTCAGGCGCCGAGACCTCATCA 186
QY	2534	CGAAGCTGTGCTGGCGGCTGACTGCGGCTGGGAGGATGGGCGAGTACCTCAAGG 2593
Db	187	CGAAGCTGTGCTGGCGGCTGACTGCGGCTGGGAGGATGGGCGAGTACCTCAAGG 246
QY	2594	CACACCGCTTCTTAACACCATCGACTTTCCCGTGACATCCGAAGCAGGCTGCACCT 2653
Db	247	CACACCGCTTCTTAACACCATCGACTTTCCCGTGACATCCGAAGCAGGCTGCACCT 306
QY	2654	ACGTCCTCCACCATCAGCCACCCCATGGACACCTCCCAATTTTGCACCGTGGATGAAGAA 2713
Db	307	ACGTCCTCCACCATCAGCCACCCCATGGACACCTCCCAATTTTGCACCGTGGATGAAGAA 366
QY	2714	GCCCTTGGCAGCGCCAGCGGAGAGCGCCAAAGCCCTGGGACACGCTGGGCTCCCCCA 2773
Db	367	GCCCTTGGCAGCGCCAGCGGAGAGCGCCAAAGCCCTGGGACACGCTGGGCTCCCCCA 426
QY	2774	GCAGCAAGCATCCAGACGAGCTTCTATGAGTTTACCTTCCGCGAGGTTCTTCGATGACA 2833
Db	427	GCAGCAAGCATCCAGACGAGCTTCTATGAGTTTACCTTCCGCGAGGTTCTTCGATGACA 486

QY 2834 ACGGCTATCCCTCCGGTGCCCGAAGCCCTCAGAGCCCGCAGAGAGTGACAGCCAGGG 2893
 Db 487 ACGGCTATCCCTCCGGTGCCCGAAGCCCTCAGAGCCCGCAGAGAGTGACAGCCAGGG 546
 QY 2894 ATGGGACTTGGAGGTGGCGGCGAGGGCTGCCAGCGGTGTACGTGTAGCCCTCAGTTA 2953
 Db 547 ATGGGACTTGGAGGTGGCGGCGAGGGCTGCCAGCGGTGTACGTGTAGCCCTCAGTTA 606
 QY 2954 ACCAACAACCTGAGGAAACCAAAATGAGATTCTTTTTCAGAGACAACTCAAGCTTAGG 3013
 Db 607 ACCAACAACCTGAGGAAACCAAAATGAGATTCTTTTTCAGAGACAACTCAAGCTTAGG 666
 QY 3014 AATCCTTCAATTTTGTCTGTTAAATGGCAACAGAGAGTCAACATGATTTCAAAT 3073
 Db 667 AATCCTTCAATTTTGTCTGTTAAATGGCAACAGAGTCAACATGATTTCAAAT 726
 QY 3074 AGCCCTCTGAGGACCTTCACTGCATTTAAACAGTATTTTAAAAAATTAGTACAGTATG 3133
 Db 727 AGCCCTCTGAGGACCTTCACTGCATTTAAACAGTATTTTAAAAAATTAGTACAGTATG 785
 QY 3134 GAAAGAGCACTTATTTGGG 3154
 Db 786 GAAAGAGCACTTATTTGTG 806

RESULT 4
 BI664597 875 bp mRNA linear EST 12-SEP-2001
 LOCUS 603290249F1 NCI_CGAP_Mam6 Mus musculus cDNA clone IMAGE:5324077 5',
 DEFINITION mRNA sequence.
 ACCESSION BI664597
 VERSION BI664597.1 GI:15578830
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE Unpublished (1999)
 JOURNAL Contact: Robert Strausberg, Ph.D.
 COMMENT Email: cgapbs-remail.nih.gov
 Tissue Procurement: Jeffrey Green M.D.
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 DNA Sequencing by: Incyte Genomics, Inc.
 Clone distribution: MGC clone distribution information can be
 found through the I.M.A.G.E. Consortium/LLNL at:
 http://image.llnl.gov
 Plate: L1AM1822 row: b column: 14
 High quality sequence stop: 798.

FEATURES
 source
 1..875
 /organism="Mus musculus"
 /strain="FVB/N"
 /db_xref="taxon:10090"
 /clone="IMAGE:5324077"
 /clone.lib="NCI_CGAP_Mam6"
 /sex="female;virgin"
 /tissue_type="infiltrating ductal carcinoma"
 /dev_stage="5 months"
 /lab_host="DH10B"
 /note="Organ: mammary; Vector: pCMV-SPORT6; Site:1: SalI;
 Site:2: NotI; Cloned unidirectionally. Primer: Oligo dt.
 Library constructed by Life Technologies, Investigator
 providing samples: Jeffrey Green, M.D., NIH"
 205 a 267 c 248 g 155 t
 BASE COUNT
 ORIGIN
 Query Match 24.5%; Score 772; DB 13; Length 875;
 Best Local Similarity 99.08; Pred. No. 3.3e-184;
 Matches 798; Conservative 0; Mismatches 5; Indels 3; Gaps 2;

QY 2128 ACTGGATTCAAGTGGAGTCAATTCACAACTACTACCAAGAAAGGAAACACATGACAG 2187
 Db 1 ACTGGATTCAAGTGGAGTCAATTCACAACTACTACCAAGAAAGGAAACACATGACAG 60
 QY 2188 GACAGATGGAGCCCGGTGACCTCTGGGACGATGTTTCCAACTGCTGCTGTGGAGACAG 2247
 Db 61 GACAGATGGAGCCCGGTGACCTCTGGGACGATGTTTCCAACTGCTGCTGTGGAGACAG 120
 QY 2248 TTAAGACCTTGGAGCAGAGGGCGCAGAGCAGCAGAGGTGCTTGGCACAATTCCTT 2307
 Db 121 TTAAGACCTTGGAGCAGAGGGCGCAGAGCAGCAGAGGTGCTTGGCACAATTCCTT 180
 QY 2308 GTCGGACACCAAAATATACATGCTCCGGAGGTCTTCCGAAAAGGTACAGCGAGCTC 2367
 Db 181 GTCGGACACCAAAATATACATGCTCCGGAGGTCTTCCGAAAAGGTACAGCGAGCTC 240
 QY 2428 TTGGCCCCACCCACACAGAGCAGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGCAT 2487
 Db 301 TTGGCCCCACCCACACAGAGCAGCAGCTGAAGGTGATCAACTGGGAGAGCAGCTGCAT 360
 QY 2488 ATCCCTACGAGTGGAGTCAAGCTGAGGCGCGAGACCTCATCAGAAAGTGTGCTGC 2547
 Db 361 ATCCCTACGAGTGGAGTCAAGCTGAGGCGCGAGACCTCATCAGAAAGTGTGCTGC 420
 QY 2548 GCGGCTGAGTCCCGCTGGGCGAGGATGGGCGAGATGACCTCAAGGCACACCCGTTCTTC 2607
 Db 421 GCGGCTGAGTCCCGCTGGGCGAGGATGGGCGAGATGACCTCAAGGCACACCCGTTCTTC 480
 QY 2608 AACACCATGACATTTTCCCGTGACATCCGAAAGCAGGCTGCACCTTACGTCCTCCACATC 2667
 Db 481 AACACCATGACATTTTCCCGTGACATCCGAAAGCAGGCTGCACCTTACGTCCTCCACATC 540
 QY 2668 AGCCACCCCTATGACACCTCCAAATTTTACCCGGTGGATGAAGAAAGCCCTGGCAGGAG 2727
 Db 541 AGCCACCCCTATGACACCTCCAAATTTTACCCGGTGGATGAAGAAAGCCCTGGCAGGAG 600
 QY 2728 GCCAGCGGAGAGCGCCCAAGGCTGGGACACGCTGGGCTTCCCTCCCGCAGCAGCATCCA 2787
 Db 601 GCCAGCGGAGAGCGCCCAAGGCTGGGACACGCTGGGCTTCCCTCCCGCAGCAGCATCCA 660
 QY 2788 GAGCAGCCCTTCTATGATGATCACTTCCGAGGTTCCTTCGATGACACAGGCTATCCCTTC 2847
 Db 661 GAGCAGCCCTTCTATGATGATCACTTCCGAGGTTCCTTCGATGACACAGGCTATCCCTTC 720
 QY 2848 CGGTGCCCGAAGCCCTCAGAGCCCGCAGAGTCCAGACCCAGGGG--ATGCGGACTTGG 2905
 Db 721 CGGTGCCCGAAGCCCTCAGAGCCCGCAGAGTCCAGACCCAGGGGGAATGCGGACTTGG 780
 QY 2906 AAGGTGCCCGCAGGCTGCCAGCG 2931
 Db 781 AAGGTGCCCG--CGAGGCTGCCAGCG 805

RESULT 5
 BM950668 737 bp mRNA linear EST 14-MAR-2002
 LOCUS UT-M-EHOP-buu-b-12-0-UI.r1 NIH_BMAP_EHOP Mus musculus cDNA clone
 DEFINITION IMAGE:5686955 5', mRNA sequence.
 ACCESSION BM950668
 VERSION BM950668.1 GI:19434258
 KEYWORDS house mouse.
 SOURCE Mus musculus
 ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 REFERENCE NIH-MGC http://mgi.nci.nih.gov/
 AUTHORS National Institutes of Health, Mammalian Gene Collection (MGC)
 TITLE

JOURNAL
COMMENT

Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library Preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

Seq primer: pYX-5.

FEATURES
source

Location/Qualifiers
1. 737
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:568695"
/clone_lib="NIH_BMAP_EH0p"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/note="Organ: brain; Vector: pYX-Asc; Site_1: Ecor I;
Site_2: Not I; The library was constructed according to
Bonald, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with Ecor I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCAGGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP). "Gene Discovery in the Developing Mouse Nervous
System", supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."

BASE COUNT
ORIGIN

Query Match 23.0%; Score 725; DB 14; Length 737;

Best Local Similarity 99.7%; Pred. No. 2.4e-172;

Matches 736; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 1370 CGAAGGGAGACAAAGCTGCGAGAGACAAAGACAGATTTCAGACCTCCCGGTGCTGTCC 1429
DB 1 CGAAGGGAGACAAAGCTGCGAGAGACAAAGACAGATTTCAGACCTCCCGGTGCTGTCC 60
QY 1430 CCAAGAAATAGCAGATGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTATG 1489
DB 61 CCAAGAAATAGCAGATGAAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTATG 120
QY 1490 CTTCAAAATTTCTATGAGCAACACGTGGAGAAATGTCATCAAAACCTTACCAGCAGAGG 1549
DB 121 CTTCAAAATTTCTATGAGC-ACACGTGGAGAAATGTCATCAAAACCTTACCAGCAGAGG 179
QY 1550 TCAGCCGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTGCTGTGAGGCCGAGC 1609
DB 180 TCAGCCGAGGCTACAGCTGGAGCAGGAAATGGCCAAAGCTGCTGTGAGGCCGAGC 239
QY 1610 AGGAGCAGATGAGGAAGATCTCTTACCAGAGGAGTCTAACTTACAAACCGCTTGAAGAGG 1669
DB 240 AGGAGCAGATGAGGAAGATCTCTTACCAGAGGAGTCTAACTTACAAACCGCTTGAAGAGG 299
QY 1670 CCAAGATGGAATGCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGG 1729
DB 300 CCAAGATGGAATGCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGG 359
QY 1730 AAGTGTGCTGCTTGTAGCTGGACATCAGCTCTGTACGCCATGAAGACTCTCAGGA 1789
DB 360 AAGTGTGCTGCTTGTAGCTGGACATCAGCTCTGTACGCCATGAAGACTCTCAGGA 419

QY 1790 AGAAGGATGCTCTGAACCGGAATCAAGTGGCCCATGCTCAAGGCTGAGAGGACATCTCG 1849
DB 420 AGAAGGATGCTCTGAACCGGAATCAAGTGGCCCATGCTCAAGGCTGAGAGGACATCTCG 479
QY 1850 CTGAAGCAGACAAATGAGTGGTGGTCAAACTCTACTTCTCTCCAGGACAAGCAGCC 1909
DB 480 CTGAAGCAGACAAATGAGTGGTGGTCAAACTCTACTTCTCTCCAGGACAAGCAGCC 539
QY 1910 TGTACTTTTGTGATGGACTACATACCAAGCGGGGATATGATGAGCCCTGCTGATCAGATGG 1969
DB 540 TGTACTTTTGTGATGGACTACATACCAAGCGGGGATATGATGAGCCCTGCTGATCAGATGG 599
QY 1970 AGGTCTTCTCTGAGCACCTGGCCCGCTTCTACATTGACAGATTGACCTGGCCATTGAAA 2029
DB 600 AGGTCTTCTCTGAGCACCTGGCCCGCTTCTACATTGACAGATTGACCTGGCCATTGAAA 659
QY 2030 GTGTCACAAAGATGGGCTTTTATCCACCGGACATCAAGCTGACAACTATCTATCAGCC 2089
DB 660 GTGTCACAAAGATGGGCTTTTATCCACCGGACATCAAGCTGACAACTATCTATCAGCC 719
QY 2090 TGGATGGTTCATTAAGC 2107
DB 720 TGGATGGTTCATTAAGC 737

RESULT 6

BB354465

LOCUS

BB354465

DEFINITION

musculus

ACCESSION

BB354465

VERSION

BB354465.2

KEYWORDS

EST

SOURCE

house mouse

ORGANISM

Mus musculus

REFERENCE

1 (bases 1 to 687)

AUTHORS

Arakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

TITLE

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

JOURNAL

Unpublished (2001)

COMMENT

On Jul 12, 2000 this sequence version replaced gi:9066293.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp

URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayashizaki,Y., Sugahara,Y., Shibata,K., Itoh

M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multipipillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a

BB354465 687 bp mRNA linear EST 24-OCT-2001
BB354465 RIKEN full-length enriched, 10 days neonate cerebellum Mus
musculus cDNA clone B930098K05 3', mRNA sequence.

BB354465
BB354465
BB354465.2 GI:16405621

EST:
Mus musculus

house mouse

Mus musculus

REFERENCE

1 (bases 1 to 687)

Akakawa,T., Carninci,P., Fukuda,S., Furuno,M., Hanagaki,T., Hara,A., Hiramoto,K., Hori,F., Ishii,Y., Ito,M., Kawai,J., Konno,H., Kouda,M., Koya,S., Matsuyama,T., Miyazaki,A., Nomura,K., Ohno,M., Okazaki,Y., Okido,T., Saito,R., Sakai,C., Sakai,K., Sano,H., Sasaki,D., Shibata,K., Shinagawa,A., Shiraki,T., Sogabe,Y., Suzuki,H., Tagami,M., Tagawa,A., Takahashi,F., Takeda,Y., Tanaka,T., Toya,T., Muramatsu,M. and Hayashizaki,Y.

RIKEN Mouse ESTs (Arakawa,T., et al. 2001)

Unpublished (2001)

On Jul 12, 2000 this sequence version replaced gi:9066293.

Contact: Yoshihide Hayashizaki

Laboratory for Genome Exploration Research Group, RIKEN Genomic

Sciences Center(GSC), Yokohama Institute

The Institute of Physical and Chemical Research (RIKEN)

1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan

Tel: 81-45-503-9222

Fax: 81-45-503-9216

Email: genome-res@sc.riken.go.jp

URL: <http://genome.gsc.riken.go.jp/>

Carninci,P., Shibata,Y., Hayashizaki,Y., Sugahara,Y., Shibata,K., Itoh

M., Konno,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y.

Normalization and subtraction of cap-trapper-selected cDNAs to

prepare full-length cDNA libraries for rapid discovery of new

genes. Genome Res. 10 (10), 1617-1630 (2000)

wagi,K., Fujiwaka,S., Inoue,K., Togawa,Y., Izawa,M., Ohara,E.,

Watanabe,M., Yoneda,Y., Ishikawa,T., Ozawa,K., Tanaka,T., Matsura

S., Kawai,J., Okazaki,Y., Muramatsu,M., Inoue,Y., Kira,A. and

Hayashizaki,Y.

RIKEN integrated sequence analysis (RISA) system--384-format

sequencing pipeline with 384 multipipillary sequencer. Genome Res.

10 (11), 1757-1771 (2000)

Konno,H., Fukunishi,Y., Shibata,K., Itoh,M., Carninci,P., Sugahara

Y. and Hayashizaki,Y.

Computer-based methods for the mouse full-length cDNA

encyclopedia: real-time sequence clustering for construction of a


```
BASE COUNT      266 a      229 c      261 g      174 t
ORIGIN
Query Match      20.2%; Score 637.4; DB 14; Length 930;
Best Local Similarity 87.4%; Pred. No. 4e-150;
Matches 721; Conservative 0; Mismatches 101; Indels 3; Gaps 2;
QY 1352 ACAAGAGCCACAAAGGTGCGAAGGAGAGCAAAAGCTGGCAGAGACAAAAGCAGATTTCAGA 1411
DB 1 ACAAGAGCCGCAAAAGCGCCRAAGGGGACAAAGCGGAAGGATATAAAGCAGATTTCAGA 60
QY 1412 CTTCCCGGCTGCTCTCCGCAAGATAGCAGAGATGAAGAGAGAGAGTCTCGCATCA 1471
DB 61 CTTCTCCCGTTCCTCCGCAAAAGCAGCAGAGAGAGAGAGAGAGAGTCAACGCATCA 120
QY 1472 AGAGTTACTCCCTTATGCTTCAATTTCTCATGGAGCAACACGCTGGAGATGTATCA 1531
DB 121 AGAGTACTCGGCATACGCTTTAAGTTCTTCATGGAGCAGCAGCTGGAGATGTATCA 180
QY 1532 AAACCTACAGCAGAGAGGTGAGCGGAGGTACAGCTGGAGCAGGAAATGGCCAAAGCTG 1591
DB 181 AAACCTACAGCAGAGAGGTTAACCGAGGCTGCAGCTGGAGCAAGAAATGGCCAAAGCTG 240
QY 1592 GCCTCTGTGAGCCGAGCAGAGAGATGAGGAAGATCTCTTACAGAGAGAGTCTTAAT 1651
DB 241 GACTCTGTGAAGCTGAGCAGAGAGAGATGCGGAAGATCTCTTACCAGAAAGAGTCTAAT 300
QY 1652 ACAACCGGCTGAGAGGCGCAGAGATGGACAAGTCCATGTTTCTGAAATCAAGACTCTAG 1711
DB 301 ACAACGAGTTAAAGAGGCGCAGAGATGGACAAGTCTATGTTTCTGAAGATCAAAACCCCTG 360
QY 1712 GCATCGGTGCTTTGGGGAAGGTGCTGCTGTGTAAGCTGGACACTACGCTCTGTAG 1771
DB 361 GGATCGGTGCTTTGGAGAGTGTGCTGTGTTGAAGTGGACACTACGCTCTGTAG 420
QY 1772 CCATGAAGACTCTCAGGAAGAGAGATGCTCTGAACCGGAATCAAGTGGCCCATGCAAG 1831
DB 421 CCATGAAGACCCCTAAGGAAAGAGATGCTCTGAACCGGAATCAGGTGGCCCATGCAAG 480
QY 1832 CTGAGAGGACATCTGCTGAGCAGACAAATGAGTGGTGTCAAACTCTACTACTCT 1891
DB 481 CCGAGAGGACATCTGCGCGGAGGAGACAAATGAGTGGTGTCAAACTCTACTACTCT 540
QY 1892 TCCAGGACAAAGACAGCCTGTACTTTGTATGAGTACATACAGCGGGGATATGATGA 1951
DB 541 TCCAGACAAAGACAGCCTGTACTTTGTATGAGTACATCTCTGTTGGGACATGATGA 600
QY 1952 GCCTCTGATCAGGATGAGGTCTTCCCTTGAGCACTGCGCCGCTTCTACATTCAGAGT 2011
DB 601 GCCTCTGATCCGGATGAGGTCTTCCCTTGAGCACTGCGCCGCTTCTACATTCGAGAGC 660
QY 2012 TGACCTGGCCATTCAAAAGTGTCCACAGATGGCTTTATCCACGGGACATCAAGCCCTG 2071
DB 661 TGACCTGGCCATTGAGAGTGTCCACAGATGGCTTTATCCACGGGACATCAAGCCCTG 720
QY 2072 ACAACATCTATCAGCTGAGTGTGATATTAAAGTGTACAGATTTTGGCCCTCTGCACT 2130
DB 721 ATAACATTTGATAGTCTGGATGTC -CATTAACATCAGATTTTGGCCCTCTGCACTG 779
QY 2131 -GGATTAGGTGGACTCAAAATTCGAAGTACTACCAAGAGGAA 2174
DB 780 GGGGTGAGGGGACCTCCCAATTCGAATATTACCCACACAGGGA 824
```

```
RESULT 8
BI331257 689 bp mRNA linear EST 30-JUL-2001
LOCUS 602982912P1 NCI_CGAP_L19 Mus musculus cDNA clone IMAGE:5135667 5'
DEFINITION mRNA sequence.
ACCESSION BI331257
VERSION BI331257.1 GI:15015901
KEYWORDS EST.
SOURCE house mouse.
```

```
ORGANISM Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 689)
AUTHORS NIH-MGC http://mgc.nci.nih.gov/.
TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
JOURNAL Unpublished (1999)
COMMENT Contact: Robert Strausberg, Ph.D.
Email: cgapbs-r@mail.nih.gov
Tissue/Procurement: Jeffrey E. Green, M.D.
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLNL1333 row: h column: 04
High quality sequence stop: 683.
Location/Qualifiers
1. 689
/organism="Mus musculus"
/strain="FVB/N"
/db_xref="taxon:10090"
/clone="IMAGE:5135667"
/lab_host="NCI_CGAP_L19"
/note="Organ: liver; Vector: pCMV-SPORT6; Site: 1; NotI;
Site: 2; SalI; Cloned unidirectionally. Primer: oligo dT.
Average insert size 1.9 kb. Constructed by Life
Technologies. Note: this is a NCI_CGAP Library."
BASE COUNT 189 a 165 c 198 g 137 t
ORIGIN
Query Match 19.7%; Score 622; DB 13; Length 689;
Best Local Similarity 97.0%; Pred. No. 2.7e-146;
Matches 656; Conservative 0; Mismatches 15; Indels 5; Gaps 2;
QY 1384 GCTGGCAGAGACAAAAGCAGATTTCAGACCTCCCGGTGCTCTCGCGCAAGATA---G 1439
DB 11 GCTGGCAGAGACAAAAGCAGATTTCAGACCTCCCGGTGCTCTCGCGCAAGATAATG 70
QY 1440 CAGAGATGAA -GAGAGAGAGAGTCTCGCATCAAGAGTTTACTCCCTTATGCTTCAAAAT 1498
DB 71 CAGAGATGAACAGAGAGAGAGTCTCGCATCAAGAGTTTACTCCCTTATGCTTCAAAAT 130
QY 1499 TCTTCATGAGCAACACGTGGAGATGTCTATCAAAACCTACCAGCAGAGGTCAGCCGA 1558
DB 131 TCTTCATGAGCAACACGTGGAGATGTCTATCAAAACCTACCAGCAGAGGTCAGCCGA 190
QY 1559 GGCTACAGCTGGAGCAGAGAAATGGCCAAAGCTGCTGTGAGGCGGAGCAGAGCAGA 1618
DB 191 GGCTACAGCTGGAGCAGAGAAATGGCCAAAGCTGCTGTGAGGCGGAGCAGAGCAGA 250
QY 1619 TGAGGAAGATCTCTTACCAGAGAGTCTTAACCTACACCGCTGGAAGAGGCGCAAGATGG 1678
DB 251 TGAGGAAGATCTCTTACCAGAGAGTCTTAACCTACACCGCTGGAAGAGGCGCAAGATGG 310
QY 1679 ACAAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGGAGTGTGCC 1738
DB 311 ACAAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGGAGTGTGCC 370
QY 1739 TCGCTTGAAGCTGGACACTCAAGCTCTGTAGCCCATGAAGACTCTCAGGAAGAAGGATG 1798
DB 371 TCGCTTGAAGCTGGACACTCAAGCTCTGTAGCCCATGAAGACTCTCAGGAAGAAGGATG 430
QY 1799 TCGTGAACCGGAATCAAGTGGCCCATGTCAAGCTGAGAGGAGACATCTCTGGTGAAGCAG 1858
DB 431 TCGTGAACCGGAATCAAGTGGCCCATGTCAAGCTGAGAGGAGACATCTCTGGTGAAGCAG 490
QY 1859 ACAATGAGTGGTGGTCAAACTCTACTACTCTCCAGCAGACAGGACGCTGTACTTTG 1918
DB 491 ACAATGAGTGGTGGTCAAACTCTACTACTCTCCAGCAGACAGGACGCTGTACTTTG 550
```


Oligo-3T track not found, Not I site shown in beginning of sequence is likely internal to the message. cDNA Library Preparation: M.B. Soares Lab Clone distribution: Researchers may obtain BMAP cDNA clones from RESEARCH GENETICS. It should be noted that Bento Soares is generating a small number of additional specialized non-redundant arrays of BMAP cDNAs whose availability will be considered under appropriate and limited collaborative arrangements The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine.
Seq primer: M13 Forward
POLYA-No.

FEATURES

Location/Qualifiers

1..575
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UI-M-CG0p-bey-c-06-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NIH_BMAP_Ret4_S2 library is a subtracted library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine
TAG_SEQ=None found"

BASE COUNT 102 a 163 c 200 g 110 t

Query Match 17.8%; Score 562.4; DB 12; Length 575;
Best Local Similarity 99.7%; Pred. No. 3e-131;
Matches 574; Conservative 0; Mismatches 1; Indels 1; Gaps 1;

QY 358 TACCACAGCTGGGTGCTCAACTACGAGGCGCCCGCGCACTGGAGAGATGCGCGG 417
DB 575 TACCACAGCTGGGTGCTCAACTACGAGGCGCCCGCGCACTGGAGAGATGCGCGG 516
QY 418 CAATATTTAGACTTCTTCTCCTGGAGCGCGGAGCCAGCCAGCGTCCAGGCTCAC 477
DB 515 CAATATTTAGACTTCTTCTCCTGGAGCGCGGAGCCAGCCAGCGTCCAGGCTCAC 457
QY 478 CAGCATCTCTCCAAAGGGTACAGCAGCAGTAGAGCAAGTGGCACTTTCCGGGCACA 537
DB 456 CAGCATCTCTCCAAAGGGTACAGCAGCAGTAGAGCAAGTGGCACTTTCCGGGCACA 397
QY 538 CACTATGGTGGTGCATCTACTATCGGAGCAGTCTGGGTATGGGTGCGAGCGCATTC 597
DB 396 CACTATGGTGGTGCATCTACTATCGGAGCAGTCTGGGTATGGGTGCGAGCGCATTC 337
QY 598 TCCTTCCAGAACAGCGCCACCATGCTTATTCAGCATGCCAAGCCAGCGGTGGC 657
DB 336 TCCTTCCAGAACAGCGCCACCATGCTTATTCAGCATGCCAAGCCAGCGGTGGC 277
QY 658 CTTCCCGGCGAGCTTCACCTTCTTCCCATGCTGGCTGTACACTGCTCGCACACAG 717
DB 276 CTTCCCGGCGAGCTTCACCTTCTTCCCATGCTGGCTGTACACTGCTCGCACACAG 217
QY 718 CCGCGGCTTACCCACCTGGGCGCCACCATATACATGTTGGGACCGCGGGTCCACAG 777
DB 216 CCGCGGCTTACCCACCTGGGCGCCACCATATACATGTTGGGACCGCGGGTCCACAG 157
QY 778 TTTACTGGGAAGCTCTGCACAGCTGTGTGGCAGCGTCCAGGAAGCGCTCAATGCT 837
DB 156 TTTACTGGGAAGCTCTGCACAGCTGTGTGGCAGCGTCCAGGAAGCGCTCAATGCT 97
QY 838 GACTTGTAGAGCTGGGTCCACAGTGGCTGTGCTGAGCTCCAGCTCCAGCGCGGAC 897
DB 96 GACTTGTAGAGCTGGGTCCACAGTGGCTGTGCTGAGCTCCAGCTCCAGCGCGGAC 37
QY 898 TCGCTGCAGAGCAGGGTCTAGAGGCTTCGCGGCGG 933

Db 36 TCGCTGCAGAGCAGGGTCTAGAGGCTTCGCGGCGG 1

RESULT 11

BF461270/c

LOCUS

DEFINITION

BF461270

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS:

TITLE

JOURNAL

MEDLINE

COMMENT

BF461270 569 bp mRNA linear EST 04-DEC-2000
UI-M-CG0p-bms-a-06-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
UI-M-CG0p-bms-a-06-0-UI 3', mRNA sequence.

BF461270.1 GI:11530432

EST.
house mouse.

Mus musculus
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.

1 (bases 1 to 569)
Normalization and subtraction: two approaches to facilitate gene
discovery

Genome Res. 6 (9), 791-806 (1996)

Contact: Chin, H
National Institute of Mental Health
6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
20992-9643, USA
Tel: 301 443 1706
Fax: 301 443 9890
Email: mEST@mail.nih.gov

Oligo-3T track not found, Not I site shown in beginning of sequence
is likely internal to the message. cDNA Library Preparation: M.B.
Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
clones from RESEARCH GENETICS. It should be noted that Bento Soares
is generating a small number of additional specialized
non-redundant arrays of BMAP cDNAs whose availability will be
considered under appropriate and limited collaborative arrangements
Seq primer: M13 Forward
POLYA-No.

Location/Qualifiers

1..569
/organism="Mus musculus"
/strain="C57BL/6J"
/db.xref="taxon:10090"
/clone="UI-M-CG0p-bms-a-06-0-UI"
/clone_lib="NIH_BMAP_Ret4_S2"
/lab_host="DH10B (Life Technologies)"
/note="Vector: pMT3D-Pac (Pharmacia) with a modified polylinker; Site.1: Not I; Site.2: Eco RI; The NIH_BMAP_Ret4_S2 library is a subtracted library, ultimately derived from mouse retina tissue libraries at various stages of development. For a detailed description of the library from which this clone was derived, please visit our web site at brainest.eng.uiowa.edu. The tissue for this library was contributed by Dr. Xin-Yuan Fu, Yale University School of Medicine
TAG_SEQ=None found"

BASE COUNT 101 a 161 c 200 g 107 t

Query Match 17.6%; Score 553.8; DB 12; Length 569;
Best Local Similarity 99.5%; Pred. No. 4.6e-129;
Matches 566; Conservative 0; Mismatches 2; Indels 1; Gaps 1;

QY 367 CTGGGTGGTGCACAACTACGAGGCGCCCGCGCACTGGAGAGATGCGCGGCAATATT 425
DB 569 CTGGGTGGTGCACAACTACGAGGCGCCCGCGCACTGGAGAGATGCGCGGCAATATT 510
QY 426 AGACTTCTCTTCCCTGGAGCGGAGCGGACCCAGCGTCCCGGCTCACACATCC 485
DB 509 AGACTTCTCTTCCCTGGAGCGGAGCGGACCCAGCGTCCCGGCTCACACATCC 450
QY 485 TCCCAAGGTTACAGCAGCAGTAGAGCAAGTGGCGGCTTCGCGGCGCACACATATGG 545

Db 449 TCCCAAGAGGTACAGCACACAGTAGAGCAAGTGGCACTTCCCGGGCACACACTATGG 390
 QY 546 TCCTGGTCACTACTATCGAGCAGTCTGGGTATGGGTGCAGCGCAGTTCCTCCTCCCA 605
 Db 389 TCCTGGTCACTACTATCGAGCAGCCTGGGTATGGGTGCAGCGCAGTTCCTCCTCCCA 330
 QY 606 GAACAAGAGCCACACAGATGCCATATCCAGCATGGCCAAAGCCAGGGTGGCCCTCCCGC 665
 Db 329 GAACAAGAGCCACACAGATGCCATATCCAGCATGGCCAAAGCCAGGGTGGCCCTCCCGC 270
 QY 566 CAGCCTCACCTTTCCTGCCCATCTGCTGGCTGTACACTGCCCTCGCACCAAGCGGGCGG 725
 Db 269 CAGCCTCACCTTTCCTGCCCATCTGCTGGCTGTACACTGCCCTCGCACCAAGCGGGCGG 210
 QY 726 TACCCACCTTGGGGCCACCACTTACATGTGTGGGCACCCGGGTGCCAGCTTACTGG 785
 Db 209 TACCCACCTTGGGGCCACCACTTACATGTGTGGGCACCCGGGTGCCAGCTTACTGG 150
 QY 786 CGAAGCTCTGCACAGCTGTGTGGCACCGCTCCAGGACAGCCTCAATGCTACTTGTGA 845
 Db 149 CGAAGCTCTGCACAGCTGTGTGGCACCGCTCCAGGACAGCCTCAATGCTACTTGTGA 90
 QY 846 CGAGCTGGCTCCACAGCTGCTGTGGCACCGCTCCAGGACAGCCTCGCTGCA 905
 Db 89 CGAGCTGGCTCCACAGCTGCTGTGGCACCGCTCCAGGACAGCCTCGCTGCA 30
 QY 906 GAAGCAGGGTCTAGAACCTTCGGGGCGGC 934
 Db 29 GAAGCAGGGTCTAGAACCTTCGGGGCGGC 1

RESULT 12

AW494954
 LOCUS
 DEFINITION UI-M-BH3-aug-e-03-0-UI.s1 NIH_BMAP_M.S4 Mus musculus cDNA clone
 UI-M-BH3-aug-e-03-0-UI 3', mRNA sequence.
 ACCESSION AW494954
 VERSION
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus

REFERENCE
 AUTHORS Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 TITLE 1 (bases 1 to 549)
 Normalization and subtraction: two approaches to facilitate gene
 discovery

JOURNAL
 MEDLINE
 COMMENT Genome Res. 6 (9), 791-806 (1996)
 9704477

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: mestr@mail.nih.gov

Oligo-dT track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 Seq primer: M13 Forward

FEATURES
 source
 POLYA-NO Location/Qualifiers
 1..549
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db_xref="taxon:10090"
 /clone="UI-M-BH3-aug-e-03-0-UI"
 /clone_lib="NIH_BMAP_M.S4"
 /dev_stage="27-32 days"

/lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site 1: Not I; Site 2: Eco RI; The
 NIH_BMAP_M.S4 library is a subtracted library of a series,
 ultimately derived from a mixture of individually tagged
 normalized libraries from ten regions of the mouse brain
 (cerebellum, brain stems, olfactory bulbs, hypothalamus,
 cortex, amygdala, basal ganglia, pineal gland, striatum,
 hippocampus) after a series of subtractions to reduce the
 representation of cDNAs from which ESTs had already been
 generated. The following serially subtracted libraries
 were generated in this process: NIH_BMAP_M.S4,
 NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, NIH_BMAP_M.S3.1,
 NIH_BMAP_M.S2, NIH_BMAP_M.S1. The subtracted library
 (NIH_BMAP_M.S4) was constructed as follows: PCR amplified
 cDNA inserts from NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and
 NIH_BMAP_M.S3.1 clones from which 3' ESTs had been derived
 was used as a driver in a hybridization with a pool of
 the NIH_BMAP_M.S3.3, NIH_BMAP_M.S3.2, and NIH_BMAP_M.S3.1
 libraries in the form of single-stranded circles. The
 remaining single-stranded circles (subtracted library)
 was purified by hydroxyapatite column chromatography,
 converted to double-stranded circles and electroporated
 into DH10B bacteria (Life Technologies) to generate the
 NIH_BMAP_M.S4 library. This procedure has been previously
 described (Bonaldio, Lennon and Soares, Genome Research
 6:791-806, 1996)
 TAG_LIB=NIH_BMAP_M.S4
 TAG_TISSUE=cerebellum
 TAG_SEQ=CACGG

BASE COUNT 126 a 181 c 167 g 73 t 2 others
 ORIGIN

Query Match 17.1%; Score 540.6; DB 10; Length 549;
 Best Local Similarity 98.9%; Pred. No. 9.8e-126;
 Matches 543; Conservative 0; Mismatches 6; Indels 0; Gaps 0;
 QY 928 CGGCCGCTATGGCTTTTCGGGCTGGCCCCCAGCAGGTACACCGCTGAGCGCGGCACACATCCTT 1047
 Db 1 CGGCCGCTATGGCTTTTCGGGCTGGCCCCCAGCAGGTACACCGCTGAGCGCGGCACACATCCTT 120
 QY 988 CTTGAGCCCTCACTGCGCGCCCCCAGCAGGTACACCGCTGAGCGCGGCACACATCCTT 1047
 Db 61 CTTGAGCCCTCACTGCGCGCCCCCAGCAGGTACACCGCTGAGCGCGGCACACATCCTT 120
 QY 1048 CACCCTGTGAAGAGCGTGGCTGTGCTGGCGCCCGAGCCCGCAGAGCGCTGGGGCCCTCG 1107
 Db 121 CACCCTGTGAAGAGCGTGGCTGTGCTGGCGCCCGAGCCCGCAGAGCGCTGGGGCCCTCG 180
 QY 1108 CACCCCGCTGGGTGGCTGGCCCCCAGCAGCTGCGCCACTGAGAGCCTGGAGAGGAG 1167
 Db 181 CACCCCGCTGGGTGGCTGGCCCCCAGCAGCTGCGCCACTGAGAGCCTGGAGAGGAG 240
 QY 1168 GGCAGCGCAGCCACACCGCTGGATGTGACTATCGCGCTCCGAGCGCAGTGGCCCA 1227
 Db 241 GGCAGCGCAGCCACACCGCTGGATGTGACTATCGCGCTCCGAGCGCAGTGGCCCA 300
 QY 1228 CCGCCTCCGTATCAAAAGCACTTGTCTGCTGCCAGTAGTCTGAGCAGTACAGCGTGGAC 1287
 Db 301 CCGCCTCCGTATCAAAAGCACTTGTCTGCTGCCAGTAGTCTGAGCAGTACAGCGTGGAC 360
 QY 1288 CTGGACAGCCTGTGACCAAGTGTGACAGAGTCTGGAGGGGGCCTGATCTAGACGGG 1347
 Db 361 CTGGACAGCCTGTGACCAAGTGTGACAGAGTCTGGAGGGGGCCTGATCTAGACGGG 420
 QY 1348 AGTCACAAGAGCCACAAAGGTGGAGGAGACAAAGCTGSCAGAGACAAAGACGATT 1407
 Db 421 AGTCACAAGAGCCACAAAGGTGGAGGAGACAAAGCTGSCAGAGACAAAGACGATT 480
 QY 1408 CAGACCTCCCGGTGCTGTGCTCCGCAAGATAGCAGATGAAGAGAGAGAGTCTCGC 1467
 Db 481 CAGACCTCCCGGTGCTGTGCTCCGCAAGATAGCAGATGAAGAGAGAGAGTCTCGC 540

QY 1468 ATCAAGAGT 1476
 Db 541 ATCAAGAGT 549

RESULT 13
 BF465780/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BF465780 539 bp mRNA linear EST 04-DEC-2000
 UI-M-CG0p-bqr-g-12-0-UI.s1 NIH_BMAP_Ret4_S2 Mus musculus cDNA clone
 UI-M-CG0p-bqr-g-12-0-UI 3', mRNA sequence.
 BF465780
 BF465780.1 GI:11534963
 EST.
 house mouse.
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 539)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 9704477

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov

Oligo-qt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 Seq primer: M13 Forward
 POLYA-No.

Location/Qualifiers
 1..539
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db.xref="taxon:10090"
 /clone="UI-M-CG0p-bqr-g-12-0-UI"
 /clone_lib="NIH_BMAP_Ret4_S2"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker. Site_1: Not I; Site_2: Eco RI; The
 NIH_BMAP_Ret4_S2 library is a subtracted library,
 ultimately derived from mouse retina tissue libraries at
 various stages of development. For a detailed description
 of the library from which this clone was derived, please
 visit our web site at brainest.eng.uiowa.edu. The tissue
 for this library was contributed by Dr. Xin-Yuan Fu, Yale
 University School of Medicine
 TAG_SEQ=None found"

97 a 151 c 189 g 102 t

BASE COUNT.
 ORIGIN

Query Match 17.0%; Score 537.4; DB 12; Length 539;
 Best Local Similarity 99.8%; Pred. No. 6.3e-125;
 Matches 538; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 396 CCACCTGGAGGAGATGCCCGCGCAATATTAGACTTCTCTCCCTGGAGCGGACCGG 455
 Db 539 CCACCTGGAGGAGATGCCCGCGCAATATTAGACTTCTCTCCCTGGAGCGGACCGG 480

QY 456 CACCCACGGTCCCGAGGCTCACAGCATCTCCCAAAGGGTACACACAGCAGTAGAGCC 515
 Db 479 CACCCACGGTCCCGAGGCTCACAGCATCTCCCAAAGGGTACACACAGCAGTAGAGCC 420

QY 516 AAGTGGCGACTTTCGGGGCACACACTATGCTGTGCTACTATCGGAGCAGTCTGG 575

Db 419 AAGTGGCGACTTTCGGGGCACACACTATGCTGTGCTACTATCGGAGCAGCTGG 360
 QY 576 GTATGGGTGCAGCGCAGTTCCTTCCAGAACAGAGCCACAGATGCTATTCCAG 635
 Db 359 GTATGGGTGCAGCGCAGTTCCTTCCAGAACAGAGCCACAGATGCTATTCCAG 300
 QY 636 CATGGCAAGGCCAGGGTGGCCCTCCCGCCAGCCTTTCCTGGCCATGCTGGGCT 695
 Db 299 CATGGCAAGGCCAGGGTGGCCCTCCCGCCAGCCTTTCCTGGCCATGCTGGGCT 240
 QY 696 GTACACTGCTCGCACCAAGCGCGGCTACCCACACCTGGGGCCACCCATTACATGT 755
 Db 239 GTACACTGCTCGCACCAAGCGCGGCTACCCACACCTGGGGCCACCCATTACATGT 180
 QY 756 GTTGGGCAACCGGGTCCACAGTTCCTTACTGGGAAAGCTTCGACAGGCTGTGTCGACCC 815
 Db 179 GTTGGGCAACCGGGTCCACAGTTCCTTACTGGGAAAGCTTCGACAGGCTGTGTCGACCC 120
 QY 816 GTCCAGGAACAGCCTCAATGCTGACTTGTACAGCTGGGCTCCACGGTGCCTGGTCTGC 875
 Db 119 GTCCAGGAACAGCCTCAATGCTGACTTGTACAGCTGGGCTCCACGGTGCCTGGTCTGC 60
 QY 876 AGCTCCACTGGCACCGCGGCTCGCTCGCAGAGCAGGCTCTAGAGCCTCGCGGCGC 934
 Db 59 AGCTCCACTGGCACCGCGGCTCGCTCGCAGAGCAGGCTCTAGAGCCTCGCGGCGC 1

RESULT 14
 AW456717/c
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

AW456717 537 bp mRNA linear EST 24-FEB-2000
 UI-M-BH3-aqv-d-11-0-UI.s1 NIH_BMAP_M_S4 Mus musculus cDNA clone
 UI-M-BH3-aqv-d-11-0-UI 3', mRNA sequence.
 AW456717
 AW456717.1 GI:7026934
 EST.
 house mouse.
 Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 537)
 Bonaldo,M.F., Lennon,G. and Soares,M.B.
 Normalization and subtraction: two approaches to facilitate gene
 discovery
 Genome Res. 6 (9), 791-806 (1996)
 9704477

Contact: Chin, H
 National Institute of Mental Health
 6001 Executive Blvd. Room 7N-7190, MSC 9643, Bethesda, MD
 20892-9643, USA
 Tel: 301 443 1706
 Fax: 301 443 9890
 Email: MEST@mail.nih.gov

Oligo-qt track not found, Not I site shown in beginning of sequence
 is likely internal to the message. cDNA Library Preparation: M.B.
 Soares Lab Clone Distribution: Researchers may obtain BMAP cDNA
 clones from RESEARCH GENETICS. It should be noted that Bento Soares
 is generating a small number of additional specialized
 non-redundant arrays of BMAP cDNAs whose availability will be
 considered under appropriate and limited collaborative arrangements
 Seq primer: M13 Forward
 POLYA-No.

Location/Qualifiers
 1..537
 /organism="Mus musculus"
 /strain="C57BL/6J"
 /db.xref="taxon:10090"
 /clone="UI-M-BH3-aqv-d-11-0-UI"
 /clone_lib="NIH_BMAP_M_S4"
 /dev_stage="27-32 days"
 /lab_host="DH10B (Life Technologies)"
 /note="Vector: pT73D-Pac (Pharmacia) with a modified
 polylinker; Site_1: Not I; Site_2: Eco RI; The

NIH_BMAP_M_S4 library is a subtracted library of a series, ultimately derived from a mixture of individually tagged normalized libraries from ten regions of the mouse brain (cerebellum, brain stems, olfactory bulbs, hypothalamus, cortex, amygdala, basal ganglia, pineal gland, striatum, hippocampus) after a series of subtractions to reduce the representation of cDNAs from which ESTs had already been generated. The following serially subtracted libraries were generated in this process: NIH_BMAP_M_S4, NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, NIH_BMAP_M_S3.1, NIH_BMAP_M_S2, NIH_BMAP_M_S1. The subtracted library (NIH_BMAP_M_S4) was constructed as follows: PCR amplified cDNA inserts from NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 clones from which 3' ESTs had been derived was used as a driver in a hybridization with a pool of the NIH_BMAP_M_S3.3, NIH_BMAP_M_S3.2, and NIH_BMAP_M_S3.1 remaining single-stranded circles (subtracted library). The library was purified by hydroxyapatite column chromatography, converted to double-stranded circles and electroporated into DH10B bacteria (Life Technologies) to generate the NIH_BMAP_M_S4 library. This procedure has been previously described (Bonaldo, Lennon and Soares, Genome Research 6:791-806, 1996)
TAG LIB=NIH_BMAP_M_S4
TAG_TISSUE=corpus-striatum
TAG_SEQ=ACGGC

BASE COUNT	98 a	151 c	186 g	102 t
ORIGIN				
Query Match	16.9%	Score 533.8;	DB 10;	Length 537;
Best Local Similarity	99.6%	Pred. No. 5.1e-124;		
Matches	535;	Conservative 0;	Mismatches 2;	Indels 0; Gaps 0;
QY	397	GCACGTGGAGGATGCGCGGCAATATTAGACTTTCTTCCTTGGAGCGGAGCGGC	456	
Db	537	GCACGTGGAGGATGCGCGGCAATATTAGACTTTCTTCCTTGGAGCGGAGCGGC	478	
QY	457	ACCCACGGTCCCGGCTCAGCAGATCTCCCAAGGTACAGCAGCAGTAGAGCCA	516	
Db	477	ACCCACGGTCCCGGCTCAGCAGATCTCCCAAGGTACAGCAGCAGTAGAGCCA	418	
QY	517	AGTGGCGCACTTTCCGGGACACACATATGGTGTGTCTACTATATCGGAGCAGTCTGGG	576	
Db	417	AGTGGCGCACTTTCCGGGACACACATATGGTGTGTCTACTATCGGAGCAGCCTGGG	358	
QY	577	TATGGGTGCAGCGCAGTTCCTCTCCAGAACAGCCACAGATGCTTATCCAGC	636	
Db	357	TATGGGTGCAGCGCAGTTCCTCTCCAGAACAGCCACAGATGCTTATCCAGC	298	
QY	637	ATGGCCAGGCGGCGGCTCCCGCCAGCCTCACCTTTCTTCCCTGCGGCTG	696	
Db	297	ATGGCCAGGCGGCGGCTCCCGCCAGCCTCACCTTTCTTCCCTGCGGCTG	238	
QY	697	TACACTGCTCGCACACAGCGGCGGTACCCACCTGGGGCCACCCATTACATGTG	756	
Db	237	TACACTGCTCGCACACAGCGGCGGTACCCACCTGGGGCCACCCATTACATGTG	178	
QY	757	TTGGGACCCCGGGTCCACGTTTACTGGCGAAAGCTGTGACAGGCTGTGTGGCACCG	816	
Db	177	TTGGGACCCCGGGTCCACGTTTACTGGCGAAAGCTGTGACAGGCTGTGTGGCACCG	118	
QY	817	TCAGAACAGCCTCAATGCTGACCTGTACAGCTGGGCTCCACGGTGGCGGTCTGCA	876	
Db	117	TCAGAACAGCCTCAATGCTGACCTGTACAGCTGGGCTCCACGGTGGCGGTCTGCA	58	
QY	877	GCTCCACTGGCAGCGGCGGCTCGTGAGAGCAGGGTCTAGAGCCTTCGCGGCGG	933	
Db	57	GCTCCACTGGCAGCGGCGGCTCGTGAGAGCAGGGTCTAGAGCCTTCGCGGCGG	1	
RESULT 15				
BE620135				

LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL
COMMENT
FEATURES
source

BE620135
601483053F1 NIH_MGC_69 Homo sapiens cDNA clone IMAGE:3805780 5',
mRNA sequence.
BE620135
BE620135.1 GI:9891073
EST.
human.
Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
1 (bases 1 to 863)
NIH-MGC http://mgc.nci.nih.gov/
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgabps-remail.nih.gov
Tissue procurement: DCTD/Dtp/Gazdar
cDNA Library Preparation: Life Technologies, Inc.
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: LLM9661 row: i column: 13
High quality sequence stop: 706.
Location/Qualifiers
1..863
/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3805780"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/note="Organ: lung; Vector: pCMV-Sport6; Site_1: NotI;
Site_2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.1 kb. Library constructed by Life
Technologies."

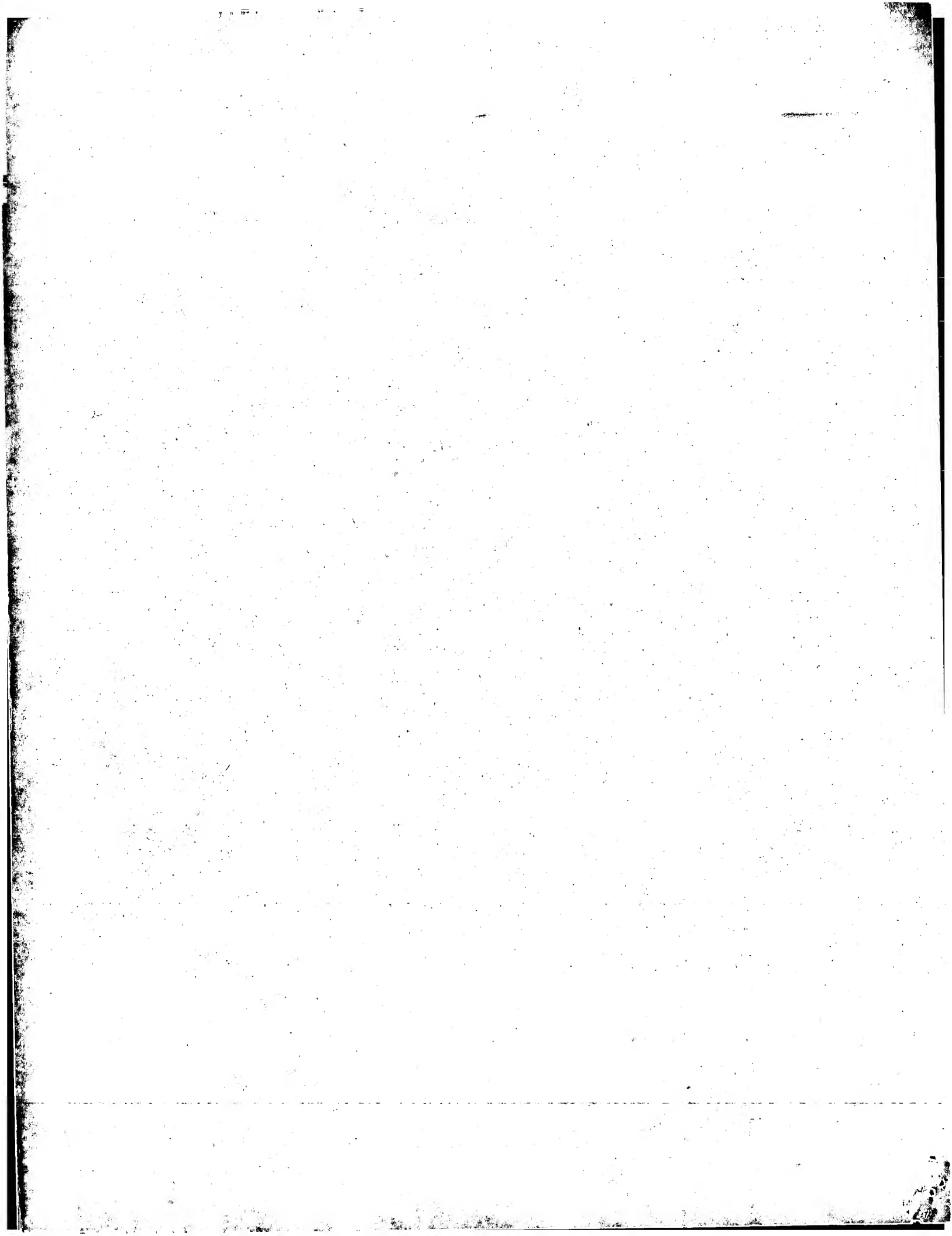
BASE COUNT	208 a	218 c	271 g	166 t
ORIGIN				
Query Match	16.5%	Score 519.4;	DB 10;	Length 863;
Best Local Similarity	88.2%	Pred. No. 2.8e-120;		
Matches	611;	Conservative 0;	Mismatches 76;	Indels 6; Gaps 4;
QY	1691	TTGTGAATCAACAGCTCTAGGATCGGTGGTGGGAGGTGCTCGCTGTGAAGC	1750	
Db	7	TTGTGAATCAACAGCTCTAGGATCGGTGGTGGGAGGTGCTCGCTGTGAAGC	65	
QY	1751	TGGACACTCAGCGCTCTAGCGCCATGAAGACTCTCAGGAAGAGGATGTCTGAACCGGA	1810	
Db	66	TGGACACTCAGCGCTCTAGCGCCATGAAGACTCTCAGGAAGAGGATGTCTGAACCGGA	125	
QY	1811	ATCAAGTGGGCGGATATGATGAGCGCTGATCAGAGTGGAGGTCTTCCCTGAGCACCTGG	1870	
Db	126	ATCAAGTGGGCGGATATGATGAGCGCTGATCAGAGTGGAGGTCTTCCCTGAGCACCTGG	185	
QY	1871	TGCTCAACTCTACTCTCTCCAGGACAGGAGCGCTTACTTGTGTGACTACATA	1930	
Db	186	TGCTCAACTCTACTCTCTCCAGGACAGGAGCGCTTACTTGTGTGACTACATA	245	
QY	1931	TACCAGCGGGGATATGATGAGCGCTGATCAGAGTGGAGGTCTTCCCTGAGCACCTGG	1990	
Db	246	TCCCTGTTGGGAGATGATGAGCGCTGATCAGAGTGGAGGTCTTCCCTGAGCACCTGG	305	
QY	1991	CCCGCTTCTACATTCAGAGTTCACCTGGCCATTGAAAGTGTCCACAAGATGGCTTTA	2050	
Db	306	CCCGCTTCTACATTCAGAGTTCACCTGGCCATTGAAAGTGTCCACAAGATGGCTTTA	365	
QY	2051	TCCACCGGACATCAAGCGCTGACACATCTATCGACCTGGATGGTATTAAGCTGA	2110	
Db	366	TCCACCGGACATCAAGCGCTGACACATCTATCGACCTGGATGGTATTAAGCTGA	425	
QY	2111	CAGATTTTGGGCTGTGCACTGGATTCAGTGGACTCACAAATCCAACTACTACCAAGAAG	2170	

```

Db 426 CAGATTTCGGCCTCTGCACTGGGTTTCAGGTGGACTCACAATTCAAAATATTACCAGAAAG 485
QY 2171 GGAACACATGAGACAGGACAGCATGGAGCCGGTGAACCTCTGGGACGATGTTTCCAACT 2230
Db 486 GGAGCCATGTCAGACAGGACAGCATGGAGCCCGGAGCCCTCTGGGATGATGTCTAACT 545
QY 2231 GTCGCTGTGGAGACAGGTTAAAGACCCCTGGAGCAGAGGGCGGAGAGCAGCAGCAGAGGT 2290
Db 546 GTCGGTGTGGGACAGGCTGAAGACCCTAGAGCAGAGGGCGGAGAGCAGCAGCAGAGGT 605
QY 2291 G--CCTGGCACATTCCTTGTGGGACACAAA--TTACATCGCTCCGGAGGTGCTTC--T 2345
Db 606 GGCCGTGGCACATTCCTTGTGGGACTCCAAACTAACATCGCACCCGAGGTGCTCCCTC 665
QY 2346 CCGCAAGGGTACACGACGCTCTGTGACTGGTG 2378
Db 666 CCGCAAGGGTACACTCAACTCTGTGACTGGGG 698

```

Search completed: January 16, 2003, 21:49:12
Job time : 2750.25 secs



GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:52 ; Search time 9497.51 Seconds
(without alignments)
17527.554 Million cell updates/sec

Title: US-09-763-334-7
Perfect score: 5720
Sequence: 1 atctagcagcgagcgagcaaa.....gcttggtgcaattagtttac 5720

Scoring table: IDENTITY_NUC
Gap 10.0, Gapext 1.0
Searched: 2054640 seqs, 14551402878 residues
Total number of hits satisfying chosen parameters: 4109280

Minimum DB seq length: 0
Maximum DB seq length: 20000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_sy.*
13: gb_un.*
14: gb_vi.*
15: em_ba.*
16: em_fun.*
17: em_in.*
18: em_in.*
19: em_mu.*
20: em_om.*
21: em_or.*
22: em_ov.*
23: em_pat.*
24: em_ph.*
25: em_pl.*
26: em_ro.*
27: em_sts.*
28: em_un.*
29: em_vi.*
30: em_htg_hum.*
31: em_htg_inv.*
32: em_htg_other.*
33: em_htg_mus.*
34: em_htg_pln.*
35: em_htg_rod.*
36: em_htg_mam.*
37: em_htg_vrt.*
38: em_sv.*
39: em_htgo_hum.*
40: em_htgo_mus.*
41: em_htgo_other.*

Pred. No. is the number of results predicted by chance to have a

score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Match	Length	DB	ID	Description
1	5708	99.8	5720	3	DMU29608	U29608 Drosophila
2	5708	99.8	5720	6	AR201454	AR201454 Sequence
3	5216.4	91.2	5360	6	AR201454	L39837 Drosophila
c	3969	69.4	163741	3	AE007821	AE007821 Drosophila
c	3942.2	68.9	236109	3	AE003775	AE003775 Drosophila
c	2965.2	51.8	28133	2	AC012975	AC012975 Drosophila
c	981.2	17.2	39839	2	AC013000	AC013000 Drosophila
8	567.8	9.9	3155	6	AR201457	AR201457 Sequence
9	565.4	9.9	3460	10	AB023958	AB023958 Mus muscu
10	558.6	9.8	5486	6	E38226	E38226 Human tumor
11	558.6	9.8	5486	6	E38227	E38227 Human tumor
12	557	9.7	3412	9	AB028019	AB028019 Homo sapi
13	557	9.7	4098	9	AF207547	AF207547 Homo sapi
14	522.8	9.1	3213	6	AR201456	AR201456 Sequence
15	522.8	9.1	3213	10	AF104414	AF104414 Mus muscu
16	470.6	8.2	3984	6	AR201455	AR201455 Sequence
17	470.6	8.2	4241	9	AF164041	AF164041 Homo sapi
18	470.6	8.2	4307	9	AF104413	AF104413 Homo sapi
19	469	8.2	2442	6	E24613	E24613 warts prote
20	455.8	8.0	1374	6	E24614	E24614 warts prote
21	307.6	5.4	181750	2	AC098623	AC098623 Rattus no
c	303.2	5.3	183597	9	AL356285	AL356285 Human DNA
23	301.2	5.3	164490	2	AC102712	AC102712 Mus muscu
24	297.4	5.2	2231	8	AF041843	AF041843 Ustilago
25	290.8	5.1	2126	3	AF238490	AF238490 Drosophila
26	290.8	5.1	2265	3	AY051880	AY051880 Drosophila
c	280.4	4.9	80146	9	AL583963	AL583963 Human DNA
27	276.4	4.8	2100	3	DMPROKINX	DMPROKINX
28	276.4	4.8	2114	3	AF239171	AF239171 Drosophila
30	262.2	4.6	215989	2	AC102227	AC102227 Mus muscu
31	261.2	4.6	2101	6	A52135	A52135 Sequence 1
32	261.2	4.6	2101	6	AR084688	AR084688 Sequence
33	256.6	4.5	3302	10	BC009658	BC009658 Mus muscu
34	254.8	4.5	2160	6	AR170896	AR170896 Sequence
35	249.6	4.4	2305	8	NCCOT1	X97657 N.crassa mR
36	239.2	4.2	245067	2	AC087136	AC087136 Mus muscu
37	232.6	4.1	3018	6	A52140	A52140 Sequence 6
38	232.6	4.1	3018	6	AR084691	AR084691 Sequence
39	232.6	4.1	3018	9	HSPROKINX	HSPROKINX
40	232.6	4.1	3593	9	BC012085	BC012085 Homo sapi
41	220.6	3.9	2202	8	MCPRTKINA	Z30329 M.crystalli
42	216.6	3.8	1935	6	AR139101	AR139101 Sequence
43	215.8	3.8	1745	3	AF275634	AF275634 Caenorhab
44	215	3.8	4725	9	BC028603	BC028603 Homo sapi
45	215	3.8	4983	6	AX056364	AX056364 Sequence

ALIGNMENTS

RESULT 1
DMU29608
LOCUS
DEFINITION Drosophila melanogaster large tumor suppressor (lats) mRNA, long
5720 bp mRNA linear INV 23-JUL-1995
transcript, complete cds.
ACCESSION U29608
VERSION U29608.1 GI:903941
KEYWORDS
SOURCE Drosophila melanogaster.
ORGANISM Drosophila melanogaster.
Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
Ephydroidea; Drosophilidae; Drosophila.
REFERENCE 1 (bases 1 to 5720)
AUTHORS Xu,T., Wang,W., Zhang,S., Stewart,R.A. and Yu,W.

FEATURES
SOURCE

gene

1
2
3
4

...

Exon

exon

COINT

Db 1021 TGTGCTGTCTGTGAGGTTAGTGTGCGAGCATTTCTGTGATGATGATGCTAAATGC 1080
 QY 1081 CACAGGGGCAAGCAGCAGCATATGATCCAGCGGCGGAAAGAGGGCGGTGCGCCCA 1140
 Db 1081 CACAGGGGCAAGCAGCAGCATATGATCCAGCGGCGGAAAGAGGGCGGTGCGCCCA 1140
 QY 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGT 1200
 Db 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGT 1200
 QY 1201 ACAAAATACATAGGAATTAATCAGAAATTTACACACCTCTGCGATACACGGGACCAACGG 1260
 Db 1201 ACAAAATACATAGGAATTAATCAGAAATTTACACACCTCTGCGATACACGGGACCAACGG 1260
 QY 1261 AGCAACGATGACATTTACTCTGACTATCACACGCGCAAGCAGCGGATGAGCGCCACC 1320
 Db 1261 AGCAACGATGACATTTACTCTGACTATCACACGCGCAAGCAGCGGATGAGCGCCACC 1320
 QY 1321 CTCGCGCTCTCTGCTCGGAGTGTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
 Db 1321 CTCGCGCTCTCTGCTCGGAGTGTATACCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1380
 QY 1381 CGAGCGGCTCCATATCCGATTCGCGTGTGGCGGTGAGTGTGGGTGTGGCGAAGCG 1440
 Db 1381 CGAGCGGCTCCATATCCGATTCGCGTGTGGCGGTGAGTGTGGGTGTGGCGAAGCG 1440
 QY 1441 ACGTGTSCCAAGATGATCAGCGCCCTTAATGCCAAACAACTGATCCGGAAGCGAGCAT 1500
 Db 1441 ACGTGTSCCAAGATGATCAGCGCCCTTAATGCCAAACAACTGATCCGGAAGCGAGCAT 1500
 QY 1501 CGAAGCGGACAGCGGAGCAGTCACTACCTGCGCTGCGCTGCGCTGCGCTGCGGAGC 1560
 Db 1501 CGAAGCGGACAGCGGAGCAGTCACTACCTGCGCTGCGCTGCGCTGCGCTGCGGAGC 1560
 QY 1561 CGGTAGTCCGATCGGAGCAGCGCCATTCGCAACACACCCACCGCGCGCTGCGAGGAC 1620
 Db 1561 CGGTAGTCCGATCGGAGCAGCGCCATTCGCAACACACCCACCGCGCGCTGCGAGGAC 1620
 QY 1621 GGTGGTAAATCAGGTGGAATGTTGGATTTCCTCGTCCGCAAGCGGTTTCAGTGAAGT 1680
 Db 1621 GGTGGTAAATCAGGTGGAATGTTGGATTTCCTCGTCCGCAAGCGGTTTCAGTGAAGT 1680
 QY 1681 GGTCCACGGCGCGCGCGCGCGCAATCCACCGCTSCAGCGGGCGCGCGCGCGCGCGCG 1740
 Db 1681 GGTCCACGGCGCGCGCGCGCGCAATCCACCGCTSCAGCGGGCGCGCGCGCGCGCGCG 1740
 QY 1741 GCGAGTCCG 1800
 Db 1741 GCGAGTCCG 1800
 QY 1801 CCGCGCGCGATGAGCG 1860
 Db 1801 CCGCGCGCGATGAGCG 1860
 QY 1861 GGTGAAGAACCG 1920
 Db 1861 GGTGAAGAACCG 1920
 QY 1921 CGGAGGCGAGTGGAGCGTGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
 Db 1921 CGGAGGCGAGTGGAGCGTGTGGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 1980
 QY 1981 AGGAGCAGCG 2040
 Db 1981 AGGAGCAGCG 2040
 QY 2041 GCGGACAAATCCCAACATCGGAGTACAGGAAATCCCGGAGCGTGGGATATCTCGCG 2100
 Db 2041 GCGGACAAATCCCAACATCGGAGTACAGGAAATCCCGGAGCGTGGGATATCTCGCG 2100
 QY 2101 CACTCGGCGGCTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 2160
 Db 2101 CACTCGGCGGCTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG

Db 2101 CACTCGGCGGCTCGCGAGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGAA 2160
 QY 2161 GCACAAACACAGTCTACAGGCCAGGAGTCAGAGCGCGCATCATCATCAGAGTGTGAA 2220
 Db 2161 GCACAAACACAGTCTACAGGCCAGGAGTCAGAGCGCGCATCATCATCAGAGTGTGAA 2220
 QY 2221 GAGCAGCAGGTCCTCAAAAGCCGTGCTGCAAAACAGCAGTGGGCGCGCAATTCGCCATCGAG 2280
 Db 2221 GAGCAGCAGGTCCTCAAAAGCCGTGCTGCAAAACAGCAGTGGGCGCGCAATTCGCCATCGAG 2280
 QY 2281 TGCCTCGGCGCGAGATTTACAGTTCACAGTGTGCGCGGTTCACACCTCTTACCTCTCAGAA 2340
 Db 2281 TGCCTCGGCGCGAGATTTACAGTTCACAGTGTGCGCGGTTCACACCTCTTACCTCTCAGAA 2340
 QY 2341 GTCGCGCGGAGTGTGTCAGCAGCAGCAACAGCAGCGCGCGCGCGCGCGCGCGCGCG 2400
 Db 2341 GTCGCGCGGAGTGTGTCAGCAGCAGCAACAGCAGCGCGCGCGCGCGCGCGCGCGCG 2400
 QY 2401 TCAGCAGCAGCAATCCAAACCCACCAACCGCCCTTGTGGTGTCTGAAYAG 2460
 Db 2401 TCAGCAGCAGCAATCCAAACCCACCAACCGCCCTTGTGGTGTCTGAAYAG 2460
 QY 2461 CAAGCCAAATTCGCTGGAGCCACCGTCTTATGCCAAGAGCATTCAGCGCGCGCGCGCG 2520
 Db 2461 CAAGCCAAATTCGCTGGAGCCACCGTCTTATGCCAAGAGCATTCAGCGCGCGCGCGCG 2520
 QY 2521 GGTGTGTAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
 Db 2521 GGTGTGTAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
 QY 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640
 Db 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2640
 QY 2641 TCAGAGGAGCGGAGTCAACCGGAGCGGAGCGGAGTCAAGAGTCTGGCGCAACGGA 2700
 Db 2641 TCAGAGGAGCGGAGTCAACCGGAGCGGAGCGGAGTCAAGAGTCTGGCGCAACGGA 2700
 QY 2701 TCTCGCGCGAGTGTCTTCGCGCGCGCGCTTATCAGAGCAACAAACAGCAGCGA 2760
 Db 2701 TCTCGCGCGAGTGTCTTCGCGCGCGCGCTTATCAGAGCAACAAACAGCAGCGA 2760
 QY 2761 GATCAAAACCGCGAGTGTCAACAAACAAACATACAGATAAAGCAACAGCAACCTTGGCGAC 2820
 Db 2761 GATCAAAACCGCGAGTGTCAACAAACAAACATACAGATAAAGCAACAGCAACCTTGGCGAC 2820
 QY 2821 GACACACCGCTTCGCGCTGCGCAATACAACTCCTCCACAGCGGCGCGAATAG 2880
 Db 2821 GACACACCGCTTCGCGCTGCGCAATACAACTCCTCCACAGCGGCGCGAATAG 2880
 QY 2881 CTCGCGCGCAGCAACGAGTCCACCGCGCACCGCTCTCTCGTCCAGCTGCAAGAA 2940
 Db 2881 CTCGCGCGCAGCAACGAGTCCACCGCGCACCGCTCTCTCGTCCAGCTGCAAGAA 2940
 QY 2941 GATCAAGCAGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 3000
 Db 2941 GATCAAGCAGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCT 3000
 QY 3001 GCGCAAGGAGTTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCG 3060
 Db 3001 GCGCAAGGAGTTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCGCTCGCG 3060
 QY 3061 GCACATAGAGAACGTGATCAAGTCTGTCGCGCGCAGCGATTCGCAAGAAATCAGCTGA 3120
 Db 3061 GCACATAGAGAACGTGATCAAGTCTGTCGCGCGCAGCGATTCGCAAGAAATCAGCTGA 3120
 QY 3121 GAAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCCAAAATCGAGATGAGGAAATGCT 3180
 Db 3121 GAAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCCAAAATCGAGATGAGGAAATGCT 3180
 QY 3181 GAACCAAAAGGAGCACTACATTCGATTAAGCGCGCGCAAGATGGCAAGAGCATGTT 3240
 Db 3181 GAACCAAAAGGAGCACTACATTCGATTAAGCGCGCGCAAGATGGCAAGAGCATGTT 3240

QY 3241 CGTCAAACTGAAGCCATTGGAGTGGGTGCAATTTGGCGAGGTAAACGCTGGTGAGCAAAAT 3300
 Db 3241 CGTCAAACTGAAGCCATTGGAGTGGGTGCAATTTGGCGAGGTAAACGCTGGTGAGCAAAAT 3300
 QY 3301 CGATACCTCGAACCAATTTGTATGGATGAACAAACCTCGGGAAGCGGAGCTTCAAGCG 3360
 Db 3301 CGATACCTCGAACCAATTTGTATGGATGAACAAACCTCGGGAAGCGGAGCTTCAAGCG 3360
 QY 3361 GAATCAGGTGCGACAGCTGAAGCGGAGAGGATATCTCGCGAAGCGGACAAATACTG 3420
 Db 3361 GAATCAGGTGCGACAGCTGAAGCGGAGAGGATATCTCGCGAAGCGGACAAATACTG 3420
 QY 3421 GGTGGTGAAGTTGACTACAGCTTCCAGGACAAGGATAATCTGTACTTTGTGATGGACTA 3480
 Db 3421 GGTGGTGAAGTTGACTACAGCTTCCAGGACAAGGATAATCTGTACTTTGTGATGGACTA 3480
 QY 3481 CATACCAAGTGGTATCTGATGTCGTCTCATCAAACTGGGCAATTTTCGAGGAGGACT 3540
 Db 3481 CATACCAAGTGGTATCTGATGTCGTCTCATCAAACTGGGCAATTTTCGAGGAGGACT 3540
 QY 3541 GGCAGATTCTACATCGCGAGGTCACCTGCGCGTGGACAGCTTCACAAATGGGCTT 3600
 Db 3541 GGCAGATTCTACATCGCGAGGTCACCTGCGCGTGGACAGCTTCACAAATGGGCTT 3600
 QY 3601 CATTCACAGACATCAAGCTGACACATATCTCATGATAGGACGACACATAAGCT 3660
 Db 3601 CATTCACAGACATCAAGCTGACACATATCTCATGATAGGACGACACATAAGCT 3660
 QY 3661 CACGACTTTGGCTGTGACGGGATTCGGATGACGACCAACTCGAAGTACTACAGGA 3720
 Db 3661 CACGACTTTGGCTGTGACGGGATTCGGATGACGACCAACTCGAAGTACTACAGGA 3720
 QY 3721 GAACGGAATCACTCGCGCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGG 3780
 Db 3721 GAACGGAATCACTCGCGCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGG 3780
 QY 3781 ACCGAAGCCACCTGTGGAGGCGACGATCGCGGATCAACAAAGAGTCTTGCGCCA 3840
 Db 3781 ACCGAAGCCACCTGTGGAGGCGACGATCGCGGATCAACAAAGAGTCTTGCGCCA 3840
 QY 3841 CTCGCTGGTGGCACCAGCTACATAGTCCCGAGGTCGTGAGAGAGTGGGTACAC 3900
 Db 3841 CTCGCTGGTGGCACCAGCTACATAGTCCCGAGGTCGTGAGAGAGTGGGTACAC 3900
 QY 3901 GCAGCTGTGCGACTACTGGAGCGTGGCGCTCATCTTAYAGATGCTGGGTGAGCC 3960
 Db 3901 GCAGCTGTGCGACTACTGGAGCGTGGCGCTCATCTTAYAGATGCTGGGTGAGCC 3960
 QY 3961 GCGCTTTCTGCGCAACAGTCCGCTGGAACGCAACAAAGGTCACTAAGTGGGAGAAAC 4020
 Db 3961 GCGCTTTCTGCGCAACAGTCCGCTGGAACGCAACAAAGGTCACTAAGTGGGAGAAAC 4020
 QY 4021 SCTGCATATTCGCGCCAGCGCGAGTTATCCGCGAGGCTACGAGCTGTAGAGGCT 4080
 Db 4021 SCTGCATATTCGCGCCAGCGCGAGTTATCCGCGAGGCTACGAGCTGTAGAGGCT 4080
 QY 4081 CTGTGCGTCTGCTGACAAAGCGGCTGGCAAGAGCTGAGAGTCAAGAGCCAGCTT 4140
 Db 4081 CTGTGCGTCTGCTGACAAAGCGGCTGGCAAGAGCTGAGAGTCAAGAGCCAGCTT 4140
 QY 4141 CTTCAAGGGCATCGACTTTTCGGGACATGCGGAAGAGCAAGAGCGCCCTACATACCGGAAT 4200
 Db 4141 CTTCAAGGGCATCGACTTTTCGGGACATGCGGAAGAGCAAGAGCGCCCTACATACCGGAAT 4200
 QY 4201 CAAGCACCACAGGACACATCTTGTATCCGCTGGATCCGGAGAGCTGGCTCGAA 4260
 Db 4201 CAAGCACCACAGGACACATCTTGTATCCGCTGGATCCGGAGAGCTGGCTCGAA 4260
 QY 4261 TGACTCCACCATGAGCAGCGGCTGATGTCGACCAAGATGACCGCAATTCACGCTT 4320
 Db 4261 TGACTCCACCATGAGCAGCGGCTGATGTCGACCAAGATGACCGCAATTCACGCTT 4320

QY 4321 TTTCCAAATTTACCTTCCTCGCTCTTTCGACGACAAAGCAGCGCGGATATGACGGAGA 4380
 Db 4321 TTTCCAAATTTACCTTCCTCGCTCTTTCGACGACAAAGCAGCGCGGATATGACGGAGA 4380
 QY 4381 TCAGGCGCGGTTTACGCTGCTGAAATGGATGCTCTCCATGTCGCCCAACACCAACACCGCGC 4440
 Db 4381 TCAGGCGCGGTTTACGCTGCTGAAATGGATGCTCTCCATGTCGCCCAACACCAACACCGCGC 4440
 QY 4441 CCCCGAATCATTTGTAGTCAAAATAGTCAAAAAGGGGATAGAAACCATTCAGTGGGCTT 4500
 Db 4441 CCCCGAATCATTTGTAGTCAAAATAGTCAAAAAGGGGATAGAAACCATTCAGTGGGCTT 4500
 QY 4501 GCATTTGAAAGGAAGCGCTGCTATAGAACTATCTATATACATATATAAATATA 4560
 Db 4501 GCATTTGAAAGGAAGCGCTGCTATAGAACTATCTATATACATATATAAATATA 4560
 QY 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATATATATATATATAT 4620
 Db 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATATATATATATATATAT 4620
 QY 4621 AT 4680
 Db 4621 AT 4680
 QY 4681 AGATGAAACGAGAGCGGAGCTAGGACCTTCGACCTTTAACTGAACATAGTATCTT 4740
 Db 4681 AGATGAAACGAGAGCGGAGCTAGGACCTTCGACCTTTAACTGAACATAGTATCTT 4740
 QY 4741 GTGCACACTACTCTCCACAAATATATATATATATATATATATATATATATAT 4800
 Db 4741 GTGCACACTACTCTCCACAAATATATATATATATATATATATATATATATAT 4800
 QY 4801 ACTGGAATGAACTTTCTGGTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 Db 4801 ACTGGAATGAACTTTCTGGTCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 QY 4861 TTAACCTAAAYGAGACGCGAATTTACCAACACTTCTCCTCTCTCTCTCTCTCTCT 4920
 Db 4861 TTAACCTAAAYGAGACGCGAATTTACCAACACTTCTCCTCTCTCTCTCTCTCTCT 4920
 QY 4921 GATCGGTGGCGGATTCGAACCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4980
 Db 4921 GATCGGTGGCGGATTCGAACCTAGAGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4980
 QY 4981 CATTCAGAAATGAGATTCGAGGTGCTGATGGAAGCAAGCAAGCAAGCAAGCAAG 5040
 Db 4981 CATTCAGAAATGAGATTCGAGGTGCTGATGGAAGCAAGCAAGCAAGCAAGCAAG 5040
 QY 5041 GGCAGCGATATAAGCGGCTCTTATAAGCTTAATCTAAATCTAACTGGGAGACAGGAC 5100
 Db 5041 GGCAGCGATATAAGCGGCTCTTATAAGCTTAATCTAAATCTAACTGGGAGACAGGAC 5100
 QY 5101 YATGATGCTGCTATCCAAATTCGCTATCTGCTATCTGCTCTCTCTCTCTCTCTCT 5160
 Db 5101 YATGATGCTGCTATCCAAATTCGCTATCTGCTATCTGCTCTCTCTCTCTCTCTCT 5160
 QY 5161 CCCCCCTCCCATCCAAAGCAAACTTAGACGTAGCTATGCTGCTGCTGCTGCTGCTGCT 5220
 Db 5161 CCCCCCTCCCATCCAAAGCAAACTTAGACGTAGCTATGCTGCTGCTGCTGCTGCTGCT 5220
 QY 5221 GTTAGACCAACTTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5280
 Db 5221 GTTAGACCAACTTTGTAATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 5280
 QY 5281 GAAATTCACACTTATCTCTGATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5340
 Db 5281 GAAATTCACACTTATCTCTGATGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 5340
 QY 5341 AAACAGTACAGAAATTTGTAATCTTCTTAAATGTAATTTGTAATTTGTAATTTGTA 5400
 Db 5341 AAACAGTACAGAAATTTGTAATCTTCTTAAATGTAATTTGTAATTTGTAATTTGTA 5400
 QY 5401 AATCTATGCTAGAGTTGCTAGCGCCCTTAAGATGTTTTTTTATGTTTATAGACCG 5460

```
Db 5401 AATCTATGCTAGAGTGTGTAGCGCCTAAGATGTTTTTTAGTTTATAGACGCTAAACGG 5460
Qy 5461 TAATCTAGTTTAAATCCCTAAACACTAAGCGAGAGTACAGTACATGCTGTTTTTTGTTGTC 5520
Db 5461 TAATCTAGTTTAAATCCCTAAACACTAAGCGAGAGTACAGTACATGCTGTTTTTTGTTGTC 5520
Qy 5521 GPAGTTCGTTGGAAATGCTTTAACGGGAACGATTTGTTTTCTCTTTAAATAGCTTCA 5580
Db 5521 GPAGTTCGTTGGAAATGCTTTAACGGGAACGATTTGTTTTCTCTTTAAATAGCTTCA 5580
Qy 5581 GTTTGTATGTCGCTGTTTTTATTTATGACTTATATATAGTCCATCTGAATTCGTGGA 5640
Db 5581 GTTTGTATGTCGCTGTTTTTATTTATGACTTATATATAGTCCATCTGAATTCGTGGA 5640
Qy 5641 TGGAGCTTATTTAAATGTGAGATCGAGCTAATTTGAAGGAAATACAAACAACTCTGTCT 5700
Db 5641 TGGAGCTTATTTAAATGTGAGATCGAGCTAATTTGAAGGAAATACAAACAACTCTGTCT 5700
Qy 5701 GCCTTGGCCAAATAGTTTAC 5720
Db 5701 GCCTTGGCCAAATAGTTTAC 5720

RESULT 2
AR201454 AR201454 5720 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 1 from patent US 6359193.
ACCESSION AR201454
VERSION AR201454.1 GI:20252342
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 5720)
AUTHORS Xu, F., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 1 19-MAR-2002;
FEATURES Location/Qualifiers
source 1. 5720
BASE COUNT 1684 a 1491 c 1457 g 1088 t
ORIGIN

Query Match 99.8%; Score 5708; DB 6; Length 5720;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 5693; Conservative 26; Mismatches 1; Indels 0; Gaps 0;

Qy 1 ATCTAGCAGCGGCGAGCAACAAACCCACGAATTAATTTACTAAATTAAGCCAAACGC 60
Db 1 ATCTAGCAGCGGCGAGCAACAAACCCACGAATTAATTTACTAAATTAAGCCAAACGC 60
Qy 61 GCATCGGAATCCCTGAAAATCGGATTAATTCACGCGGAAAGTATGATGGTTGCGAAGC 120
Db 61 GCATCGGAATCCCTGAAAATCGGATTAATTCACGCGGAAAGTATGATGGTTGCGAAGC 120
Qy 121 GAGTGAATCAAGTGAATATAGTTCGGCAATATACGCAATTTGCTCAAAAGGCAAGGA 180
Db 121 GAGTGAATCAAGTGAATATAGTTCGGCAATATACGCAATTTGCTCAAAAGGCAAGGA 180
Qy 181 AAACGGAGAAAAGAGGAAAAGCAATTAAGTCCGCTGTGGGAAACGCGAAGGCGGA 240
Db 181 AAACGGAGAAAAGAGGAAAAGCAATTAAGTCCGCTGTGGGAAACGCGAAGGCGGA 240
Qy 241 GAACAAAGCGGAAACGAGGAAATTCGCTGGAACGCGGAAACGCGAAGGCGGA 300
Db 241 GAACAAAGCGGAAACGAGGAAATTCGCTGGAACGCGGAAACGCGAAGGCGGA 300
Qy 301 AGCTCCAAAGTTGGCCGCCCATCGATTTCGTGAGGATCAATTAAGATTCCGAGTGGTCA 360
Db 301 AGCTCCAAAGTTGGCCGCCCATCGATTTCGTGAGGATCAATTAAGATTCCGAGTGGTCA 360
Qy 361 GAATCGGCTCAAAATCAAAATAAAATCAAAATTAATTTGGTATTCAGATATTCAAATGGA 420
```

```
Db 361 GAATCGGCTCAAAATCAAAATAAAATCAAAATTAATTTGGTATTCAGATATTCAAATGGA 420
Qy 421 ATTCAATTCATCCCTGCGGACTTTTATTCCGAGTCTGCCAACTATTTTGAATTTGAATGT 480
Db 421 ATTCAATTCATCCCTGCGGACTTTTATTCCGAGTCTGCCAACTATTTTGAATTTGAATGT 480
Qy 481 GTGCTGCGGCTGCGGCGAGAAATCTCTGATAAAGAGAGAGGAATAAAATCGGAAGAACAA 540
Db 481 GTGCTGCGGCTGCGGCGAGAAATCTCTGATAAAGAGAGAGGAATAAAATCGGAAGAACAA 540
Qy 541 AATACAAATACAAATGAAATGCGGGAGCAGATTTTACATGCCAAATGAATCTGATAG 600
Db 541 AATACAAATACAAATGAAATGCGGGAGCAGATTTTACATGCCAAATGAATCTGATAG 600
Qy 601 GCGAAGGGGGGTTCTCTTATTAATGCAAAATGTAATGTAATGTAATGTAATGTAATGTAAT 660
Db 601 GCGAAGGGGGGTTCTCTTATTAATGCAAAATGTAATGTAATGTAATGTAATGTAATGTAAT 660
Qy 661 GTGGAAGAAATCCCGCGCGAGTGATTAATAATCGGACGACAAACAAACAGCAGAGCCTAC 720
Db 661 GTGGAAGAAATCCCGCGCGAGTGATTAATAATCGGACGACAAACAAACAGCAGAGCCTAC 720
Qy 721 ACCGCGAGAAAGAGCAGCGCAACAAATATCTTTTATTTAGAGCAACAAATATCAAGATC 780
Db 721 ACCGCGAGAAAGAGCAGCGCAACAAATATCTTTTATTTAGAGCAACAAATATCAAGATC 780
Qy 781 GAGATAATAAAGCATCTTAAACCCGCGCTTAGTTCGTTTGTAGTCTGCCACRATATA 840
Db 781 GAGATAATAAAGCATCTTAAACCCGCGCTTAGTTCGTTTGTAGTCTGCCACRATATA 840
Qy 841 GATATTCAAAAGCAAAAGGTGTGTGCGCATCGCAGACAAACAAACAGTAAAGCATCTATT 900
Db 841 GATATTCAAAAGCAAAAGGTGTGTGCGCATCGCAGACAAACAAACAGTAAAGCATCTATT 900
Qy 901 TCATACAAACAAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 960
Db 901 TCATACAAACAAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATA 960
Qy 961 CAATCAAAATCCCGCGCGCGCATGTGCCCCAGTGTGCGGTGTGCGGTGTGCGGTGTGCG 1020
Db 961 CAATCAAAATCCCGCGCGCGCATGTGCCCCAGTGTGCGGTGTGCGGTGTGCGGTGTGCG 1020
Qy 1021 TGTGCTGTGCTGTGCGAGTGTGTGCGGAGCATTTCTGTGATATGATGCTCTAAATGC 1080
Db 1021 TGTGCTGTGCTGTGCGAGTGTGTGCGGAGCATTTCTGTGATATGATGCTCTAAATGC 1080
Qy 1081 CACAGGGGGAAGCAGCAGCATCATGCCAGGGGCGGAAAAAGGGGGGGTGCAGCCCAA 1140
Db 1081 CACAGGGGGAAGCAGCAGCATCATGCCAGGGGCGGAAAAAGGGGGGGTGCAGCCCAA 1140
Qy 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTAAACCCGATTTGAAGT 1200
Db 1141 TGATAAATACAGCGGGAAGCCCTCGAGAGCATCAAGCAGGACCTAAACCCGATTTGAAGT 1200
Qy 1201 ACAAATAACCATAGGAATAATACAGAAATTCACACTCTCGCATACACGGCGACCAACGG 1260
Db 1201 ACAAATAACCATAGGAATAATACAGAAATTCACACTCTCGCATACACGGCGACCAACGG 1260
Qy 1261 AGGCAAGCATGACATCTCTCTGACTATCACCCAGCAGCAGCGGATGGAGCCGCCACC 1320
Db 1261 AGGCAAGCATGACATCTCTCTGACTATCACCCAGCAGCAGCGGATGGAGCCGCCACC 1320
Qy 1321 CTCGCCCTCTCTGCTCGGAGTGTGTATACCCGCCGCCGCCCATTTAGGTCAGCC 1380
Db 1321 CTCGCCCTCTCTGCTCGGAGTGTGTATACCCGCCGCCGCCCATTTAGGTCAGCC 1380
Qy 1381 CGGAGCCGGCTCCATATCCGTATCCGTTGTGGCGCTTGGAGTGGTGGTGGTGGTGGTGGT 1440
Db 1381 CGGAGCCGGCTCCATATCCGTATCCGTTGTGGCGCTTGGAGTGGTGGTGGTGGTGGTGGT 1440
Qy 1441 AGTGTSCCAAGATGATGACGGCCCTAATGCCAAACAACTATCCGGAAGCCGAGCAT 1500
```

Db	1441	ACGTGTC	CAAGA	TGATG	ACGGCC	CTAATG	CAAA	AAACTG	ATCGG	AAACG	CGAGCAT	1500
Qy	1501	CGA	CGGAC	AGCG	GAG	CAGT	CTAC	TCTCG	CTG	CGCTCG	CTG	1560
Db	1501	CGA	CGGAC	ACG	CGC	GAG	CTAC	TCTCG	CTG	CGCTCG	CTG	1560
Qy	1561	CGT	AGCTCC	GATCG	ACG	CCCC	CAATTC	CG	ACCA	CAACCC	ACG	1620
Db	1561	CGT	AGCTCC	GATCG	ACG	CCCC	CAATTC	CG	ACCA	CAACCC	ACG	1620
Qy	1621	GGT	GGGTAA	TCCAG	GTG	GAATTC	CGT	CGC	CAAG	CGGTTT	CAGT	1680
Db	1621	GGT	GGGTAA	TCCAG	GTG	GAATTC	CGT	CGC	CAAG	CGGTTT	CAGT	1680
Qy	1691	GGT	CCAC	CGCG	CGCC	GCAC	AGCA	TCCAC	CGCT	SCAG	CGCG	1740
Db	1691	GGT	CCAC	CGCG	CGCC	GCAC	AGCA	TCCAC	CGCT	SCAG	CGCG	1740
Qy	1741	GCC	AGTCC	CGCC	AC	AGCC	AGG	CGGTAG	TTGA	AGCG	CGG	1800
Db	1741	GCC	AGTCC	CGCC	AC	AGCC	AGG	CGGTAG	TTGA	AGCG	CGG	1800
Qy	1801	CCG	CGCG	CGAT	AG	CGCC	AC	CTAC	AG	CGMG	CAACT	1860
Db	1801	CCG	CGCG	CGAT	AG	CGCC	AC	CTAC	AG	CGMG	CAACT	1860
Qy	1861	GCT	GAAG	AAC	CGC	AGC	AGC	AGT	GA	CGC	AGC	1920
Db	1861	GCT	GAAG	AAC	CGC	AGC	AGC	AGT	GA	CGC	AGC	1920
Qy	1921	CGG	AGC	AGT	GG	AGC	AGT	GG	AGC	CGC	CTAC	1980
Db	1921	CGG	AGC	AGT	GG	AGC	AGT	GG	AGC	CGC	CTAC	1980
Qy	1981	AGG	AGC	AGC	CGC	CGC	CGC	CGC	CGC	CGC	CGC	2040
Db	1981	AGG	AGC	AGC	CGC	CGC	CGC	CGC	CGC	CGC	CGC	2040
Qy	2041	GCC	AC	CA	CAAT	CC	CA	CAAT	CC	CA	CAAT	2100
Db	2041	GCC	AC	CA	CAAT	CC	CA	CAAT	CC	CA	CAAT	2100
Qy	2101	CAC	CTCG	GGG	GGT	CGC	CG	AGC	CGC	CGC	CGC	2160
Db	2101	CAC	CTCG	GGG	GGT	CGC	CG	AGC	CGC	CGC	CGC	2160
Qy	2161	GCC	AC	AC	AGC	AGT	CTAC	CG	CGC	CGC	CGC	2220
Db	2161	GCC	AC	AC	AGC	AGT	CTAC	CG	CGC	CGC	CGC	2220
Qy	2221	GAG	CAC	CG	AGT	CC	AA	AGC	CGC	CGT	GTG	2280
Db	2221	GAG	CAC	CG	AGT	CC	AA	AGC	CGC	CGT	GTG	2280
Qy	2281	TGC	CTCG	CGC	AG	CAAT	CC	AGT	CC	AGT	CC	2340
Db	2281	TGC	CTCG	CGC	AG	CAAT	CC	AGT	CC	AGT	CC	2340
Qy	2341	GT	CCG	CGC	AGT	GTG	CA	AGC	AGC	AGC	AGC	2400
Db	2341	GT	CCG	CGC	AGT	GTG	CA	AGC	AGC	AGC	AGC	2400
Qy	2401	TC	AGC	AC	AGC	AA	TC	CA	AGC	AA	TC	2460
Db	2401	TC	AGC	AC	AGC	AA	TC	CA	AGC	AA	TC	2460
Qy	2461	CA	AGC	CA	AA	TTC	CGT	CGC	CGC	CGC	CGC	2520
Db	2461	CA	AGC	CA	AA	TTC	CGT	CGC	CGC	CGC	CGC	2520
Qy	2521	GGT	GTT	AT	AGC	AGC	AGC	AGC	AGC	AGC	AGC	2580
Db	2521	GGT	GTT	AT	AGC	AGC	AGC	AGC	AGC	AGC	AGC	2580

QY	2581	ACAGCAGCAGCAGCAGCAGCAGCAACTGCAGGCCTTGAGGGTGTCTCCAGGCACAGGC	2640
DB	2581		
QY	2581	ACAGCAGCAGCAGCAGCAGCAGCAACTGCAGGCCTTGAGGGTGTCTCCAGGCACAGGC	2640
DB	2581		
QY	2641	TCAGAGGAGCGGGATCAACGGGAGCGGGAACGGGATCAGCAGAGCTGGCCACACGAAA	2700
DB	2641		
QY	2701	TCCTGGCGCGCAGATGCTCCGCGCGGCCCTATCAGAGCAACAACAACAACAGCGA	2760
DB	2701		
QY	2701	TCCTGGCGCGCAGATGCTCCGCGCGGCCCTATCAGAGCAACAACAACAACAGCGA	2760
DB	2701		
QY	2761	GATCAAAACGCGCAGCTGCAACAACAACAACATACAGATACAGCAACAGCACTGGCCAC	2820
DB	2761		
QY	2821	GACACCAACCTTCGGCTGCCAAATAAATAAATACTCTCAACACAGCGCGCGCAATAG	2880
DB	2821		
QY	2821	GACACCAACCTTCGGCTGCCAAATAAATAAATACTCTCAACACAGCGCGCGCAATAG	2880
DB	2821		
QY	2881	CTCGGCGCGCAGCAACCGATCCACGGCACACACGGCTCTCTGCACACAGCTGCAAGAA	2940
DB	2881		
QY	2941	GATCAAGCAGCCTCGCCATCCGGAGCGCAAGAATCTCCAAGGAAGAGGAGGAGA	3000
DB	2941		
QY	3001	GCSCAAGGATGCCGCATCAGCAGTACTGCCGCCAAGCCCTCAAGTCTTCATGAGCA	3060
DB	3001		
QY	3061	GCACATAGAGAGTGATCAAGTCTATCGCCAGCGCAGCTATCGCAAGAATCAGCTGGA	3120
DB	3061		
QY	3121	GAAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAATGCT	3180
DB	3121		
QY	3181	GAACCAAAAGGAGAGCAACTACATTCGATTAAGCGCGCAAGATGGACAAGACATGTT	3240
DB	3181		
QY	3241	CGTCAAACTGAAGCCCATTTGGATGGGTGCAATTTGGCAGGTAAAGCTGGTGCACAAAT	3300
DB	3241		
QY	3301	CGATACCTCGAACCATTTGATCGGATGAACCCCTCGCGAAGCGGACGTTCTCAAGCG	3360
DB	3301		
QY	3361	GAATCAGTGGCACACGTGAAGCGGAGGGGATCTCTCGCGAAGCGCGACATAACTG	3420
DB	3361		
QY	3421	GGTGTGAAGTTGTATACAGCTCCAGGACAGGATATCTGTACTTTGTGATGACTA	3480
DB	3421		
QY	3481	CATACCGAGTGGTATCTGATGCTCGCTCATCAAACTGGCATTTTCGAGGAGAACT	3540
DB	3481		
QY	3541	GGCCAGATTCTACATCGCGGAGGTCACTGCGCGTGGACAGCGTTTCACAAAATGGCTT	3600
DB	3541		
QY	3601	CATTACAGACATCAAGCTTGACAACATACTCATGATAGGACGGACACATAAAGCT	3660
DB	3601		

QY 3661 CACCGACTTTGGCTGTGTGACGGGATTCGATGGAGCCCAAACTCGAAGTACTTACCAGGA 3720
 Db 3661 CACCGACTTTGGCTGTGTGACGGGATTCGATGGAGCCCAAACTCGAAGTACTTACCAGGA 3720
 QY 3721 GAACGGCAATCACTCCGCGCAGGACTCGATGAGCCCTGGGAGGAAATCTCCGAGAACGG 3780
 Db 3721 GAACGGCAATCACTCCGCGCAGGACTCGATGAGCCCTGGGAGGAAATCTCCGAGAACGG 3780
 QY 3781 ACCGAAGCCACCGTCTGCGAGAGGCGGAGTCCGCGATCAACAAAGAGTCTCTGCCCA 3840
 Db 3781 ACCGAAGCCACCGTCTGCGAGAGGCGGAGTCCGCGATCAACAAAGAGTCTCTGCCCA 3840
 QY 3841 CTCGCTGGTGGGACCCGAACTACATAGCTCCCGAGGCTGCTGGAGAGTGGGTACAC 3900
 Db 3841 CTCGCTGGTGGGACCCGAACTACATAGCTCCCGAGGCTGCTGGAGAGTGGGTACAC 3900
 QY 3901 GCAGCTGTCGGACTACTGAGCGGTGGCGTCACTCTTAYAGATGCTGGTGGGTACGC 3960
 Db 3901 GCAGCTGTCGGACTACTGAGCGGTGGCGTCACTCTTAYAGATGCTGGTGGGTACGC 3960
 QY 3961 GCCCTTTCTGGCCAAAGTCCGCTGGAAACGCAAAAGGTCATCAACTGGGAGAAAC 4020
 Db 3961 GCCCTTTCTGGCCAAAGTCCGCTGGAAACGCAAAAGGTCATCAACTGGGAGAAAC 4020
 QY 4021 SCTGCATATTCGCGCGCAGCGGAGTTATCCCGGAGGCTACGACTTTGATAAGGAGCT 4080
 Db 4021 SCTGCATATTCGCGCGCAGCGGAGTTATCCCGGAGGCTACGACTTTGATAAGGAGCT 4080
 QY 4081 CTGTGCGTGGTGCACAGCGGTGGGCAAGAGCGTGGAGAGGTCAGAGCCACGACTT 4140
 Db 4081 CTGTGCGTGGTGCACAGCGGTGGGCAAGAGCGTGGAGAGGTCAGAGCCACGACTT 4140
 QY 4141 CTTCAAGGCGCATCGACTTTTGGGACATCGGAAGCAGAAAGCGCCCTACATACCGGAAT 4200
 Db 4141 CTTCAAGGCGCATCGACTTTTGGGACATCGGAAGCAGAAAGCGCCCTACATACCGGAAT 4200
 QY 4201 CAAGCACCCACGAGGCGGATGATGTCGACCAAGATGACCGGCTTCCACGGCTT 4320
 Db 4201 CAAGCACCCACGAGGCGGATGATGTCGACCAAGATGACCGGCTTCCACGGCTT 4320
 QY 4321 TTTTGAATTTACTCTCCGCTGCTTCTGACGACAGAGCGCCGGATATGACGGAGCA 4380
 Db 4321 TTTTGAATTTACTCTCCGCTGCTTCTGACGACAGAGCGCCGGATATGACGGAGCA 4380
 QY 4381 TCAGGCGCGGTTTACGCTGAAATGGATGCTTCCATGTGCCCAACACACACCCCGC 4440
 Db 4381 TCAGGCGCGGTTTACGCTGAAATGGATGCTTCCATGTGCCCAACACACACCCCGC 4440
 QY 4441 CCCCAGATCATTTAGTCAATAGTCAAAAAGGGATAGAAACCATGAGTGGGCTT 4500
 Db 4441 CCCCAGATCATTTAGTCAATAGTCAAAAAGGGATAGAAACCATGAGTGGGCTT 4500
 QY 4501 GCATTGTAAGGAAGCGTGGCTATAGAATGAACATCTATATACATTAATAAATATA 4560
 Db 4501 GCATTGTAAGGAAGCGTGGCTATAGAATGAACATCTATATACATTAATAAATATA 4560
 QY 4561 GGAGACAGTGGAGGGGAGCTACGCTATACATCAAAAATATATATATATTTGATAT 4620
 Db 4561 GGAGACAGTGGAGGGGAGCTACGCTATATACATCAAAAATATATATATATTTGATAT 4620
 QY 4621 ATATATATATATATGCGGTAGGCGATGAATGAATAATAATAACGAGCCGAGTAG 4680
 Db 4621 ATATATATATATATGCGGTAGGCGATGAATGAATAATAATAACGAGCCGAGTAG 4680
 QY 4681 AGATGAACGAGAGGCGGAGTCCGACCTTCGACCTTTAACTGACATAGTATATCTT 4740
 Db 4681 AGATGAACGAGAGGCGGAGTCCGACCTTCGACCTTTAACTGACATAGTATATCTT 4740
 QY 4741 GTGCACTACTACTCCCAACAAATATATATTTTAAATTTGTAGATTTCAAAAGGGRCCA 4800

Db 4741 GTGCACTACTACTCCCAACAAATATATATTTTAAATTTGTAGATTTCAAAAGGAGCA 4800
 QY 4801 ACTGGAAATCGAACCTTTCTGTGTCTCAAGCAAGCAAGCAAAACAAACAGCC 4860
 Db 4801 ACTGGAAATCGAACCTTTCTGTGTCTCAAGCAAGCAAGCAAAACAAACAGCC 4860
 QY 4861 TTAACATAAAGAGAGCGGAAATTTACCAACACATTCCTCTCTCTCTCTCTCTCTCT 4920
 Db 4861 TTAACATAAAGAGAGCGGAAATTTACCAACACATTCCTCTCTCTCTCTCTCTCTCT 4920
 QY 4921 GATCGGTGCCGATTCGAACCTAGCAGAGCTGGTTCATCCGCGCCATCCATTTACTTCC 4980
 Db 4921 GATCGGTGCCGATTCGAACCTAGCAGAGCTGGTTCATCCGCGCCATCCATTTACTTCC 4980
 QY 4981 CATTCAGAAATTCAGATTCGAGGTGGCGATGGAGAACGAGGAGACCAAAAGTCGCAC 5040
 Db 4981 CATTCAGAAATTCAGATTCGAGGTGGCGATGGAGAACGAGGAGACCAAAAGTCGCAC 5040
 QY 5041 GGCAGCGATATAGCGGGTCTTATAGCCCTAATCTAAATCTAACTGGGAGAACAGGACC 5100
 Db 5041 GGCAGCGATATAGCGGGTCTTATAGCCCTAATCTAAATCTAACTGGGAGAACAGGACC 5100
 QY 5101 YATGATGCTGCTATCCAAATTCGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
 Db 5101 YATGATGCTGCTATCCAAATTCGCTATCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 5160
 QY 5161 CCCCCCTTCCCTCCCAACAAAGCAAACTTAGACGTAGCTATGTGAAAAGCTAGCAAT 5220
 Db 5161 CCCCCCTTCCCTCCCAACAAAGCAAACTTAGACGTAGCTATGTGAAAAGCTAGCAAT 5220
 QY 5221 GTTAGACCAACTTGTGAATGCAATGAAATGTTTACGCCACGAGGAAACGCGGG 5280
 Db 5221 GTTAGACCAACTTGTGAATGCAATGAAATGTTTACGCCACGAGGAAACGCGGG 5280
 QY 5281 GAAATTCACACTTATCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
 Db 5281 GAAATTCACACTTATCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
 QY 5341 AAACAGTACGAGAAATTCATCTCTTAATGTAATATGTAAGAACACACGTTAATGT 5400
 Db 5341 AAACAGTACGAGAAATTCATCTCTTAATGTAATATGTAAGAACACACGTTAATGT 5400
 QY 5401 AATCTATGCTAGAGTGTGTAGCGCTAAGATGTTTTTATGTTTATAGACGCTAACCG 5460
 Db 5401 AATCTATGCTAGAGTGTGTAGCGCTAAGATGTTTTTATGTTTATAGACGCTAACCG 5460
 QY 5461 TAATCTAGTTTAAATTCCTAACACTAAGCGAGAGTACAGTACATGTTTTTTTGTGTC 5520
 Db 5461 TAATCTAGTTTAAATTCCTAACACTAAGCGAGAGTACAGTACATGTTTTTTTGTGTC 5520
 QY 5521 GTAGGTTCCTTGGAATGCTTAAACGGGAAACGATTTGTTTCTCTTTAATAGCTTCA 5580
 Db 5521 GTAGGTTCCTTGGAATGCTTAAACGGGAAACGATTTGTTTCTCTTTAATAGCTTCA 5580
 QY 5581 GTTTGATGCTGCTGTTTTTATATGACTTATATATATATAGTCCATCTGAATATTCGTGA 5640
 Db 5581 GTTTGATGCTGCTGTTTTTATATGACTTATATATATATAGTCCATCTGAATATTCGTGA 5640
 QY 5641 TGGAGCCTATTTAAATGTGAGATCGAGTAAATGAAGGAAATACAAACAACTCTGTCT 5700
 Db 5641 TGGAGCCTATTTAAATGTGAGATCGAGTAAATGAAGGAAATACAAACAACTCTGTCT 5700
 QY 5701 GCCTTGGCCAAATAGTTTAC 5720
 Db 5701 GCCTTGGCCAAATAGTTTAC 5720

RESULT 3

DROWARTS

LOCUS

DEFINITION

Drosophila melanogaster tumor suppressor (warts) mRNA linear INV 03-AUG-1995
 complete cds.

us-09-763-334-7.rge

[illegible]

QY	3119	GAGAAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAAATG	31178
Db	3153	GAGAAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAAATG	3212
QY	3179	CTGAACCAAAAGGAGAGCAACTACATCTCGAATCGAAGCGGCCAAAGATGACAAGAGCATG	3238
Db	3213	CTGAACCAAAAGGAGAGCAACTACATCTCGAATCGAAGCGGCCCAAGATGACAAGAGCATG	3272
QY	3239	TTCTGTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTAACTGCTGTGAGCAAA	3298
Db	3273	TTCTGTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTAACTGCTGTGAGCAAA	3332
QY	3299	ATCGATACCTCGAACCATTTTGTATGCCATGAATAACCTCGGGAAGCGGACGCTTCTCAAG	3358
Db	3333	ATCGATACCTCGAACCATTTTGTATGCCATGAATAACCTCGGGAAGCGGACGCTTCTCAAG	3392
QY	3359	CGGAATCAGGTGGCACACGCTGAAGCCCGAGAGGGATATCCTCGCGAAGCCGACAAATAAC	3418
Db	3393	CGGAATCAGGTGGCACACGCTGAAGCCCGAGAGGGATATCCTCGCGAAGCCGACAAATAAC	3452
QY	3419	TGGGTGGTGAAGTTTGTACTACAGCTTCCAGGACAAGGATAACTCTGATCTTGTGATGGAC	3478
Db	3453	TGGGTGGTGAAGTTTGTACTACAGCTTCCAGGACAAGGATAACTCTGATCTTGTGATGGAC	3512
QY	3479	TACATACCAAGTGGTGATCTGATCTGCTGCTCATCAAACTGGGCACTTTTCGAGGAGGAA	3538
Db	3513	TACATACCAAGTGGTGATCTGATCTGCTGCTCATCAAACTGGGCACTTTTCGAGGAGGAA	3572
QY	3539	CTGGCCAGATTTACATCCGCCGAGGTACCTGCGCCCGTGGACAGCGTTTACAAAATGGCC	3598
Db	3573	CTGGCCAGATTTACATCCGCCGAGGTACCTGCGCCCGTGGACAGCGTTTACAAAATGGCC	3632
QY	3599	TTCAATTCACAGAGACATCAAGCTTGACAACTACTCATAGGACGCGACACATATAAG	3658
Db	3633	TTCAATTCACAGAGACATCAAGCTTGACAACTACTCATAGGACGCGACACATATAAG	3692
QY	3659	CTCACCGACTTTTGGCTGTGCACGGGATTCGATGGACGCGACAACTCGAAGTACTACCAAG	3718
Db	3693	CTCACCGACTTTTGGCTGTGCACGGGATTCGATGGACGCGACAACTCGAAGTACTACCAAG	3752
QY	3719	GAGAAGCGCAATCACTCGCCGACGACNCTGATGAGCCCTGGGAGGAATACTCCGAGAAC	3778
Db	3753	GAGAAGCGCAATCACTCGCCGACGACNCTGATGAGCCCTGGGAGGAATACTCCGAGAAC	3812
QY	3779	GGACCGAAGCCCAACCGTGTGTGAGAGCGGACGGATGCGCGATCACCAAGAGTCTCTGGCC	3838
Db	3813	GGACCGAAGCCCAACCGTGTGTGAGAGCGGACGGATGCGCGATCACCAAGAGTCTCTGGCC	3872
QY	3839	CATCTCGTGTGGGCACCCCGAACTACATAGCTCCGAGGTGCTGGAGAGTCTGGGTAC	3898
Db	3873	CATCTCGTGTGGGCACCCCGAACTACATAGCTCCGAGGTGCTGGAGAGTCTGGGTAC	3932
QY	3899	ACGCAGCTGTGCGACTACTGGAGCGTGGCGGCTCATCTCTTAYGAGATGCTGTGGGTTCAG	3958
Db	3933	ACGCAGCTGTGCGACTACTGGAGCGTGGCGGCTCATCTCTTAYGAGATGCTGTGGGTTCAG	3992
QY	3959	CCGCGCTTTCTGGCCAAACAGTCCGCTGGGAAACGCGCAAAAAGGTCTAACTGGGAGAAA	4018
Db	3993	CCGCGCTTTCTGGCCAAACAGTCCGCTGGGAAACGCGCAAAAAGGTCTAACTGGGAGAAA	4052
QY	4019	ACSTCTGATATTCGCGCGACGCGGAGTTATCCGCGAGGCTACGGACTTGTATAAGGAGG	4078
Db	4053	ACSTCTGATATTCGCGCGACGCGGAGTTATCCGCGAGGCTACGGACTTGTATAAGGAGG	4112
QY	4079	CTCTGTGCGTCTGGCTGACAAAGCGGCTGGGCAAGAGCGTGGACGAGTCAAGACCGACAC	4138
Db	4113	CTCTGTGCGTCTGGCTGACAAAGCGGCTGGGCAAGAGCGTGGACGAGTCAAGACCGACAC	4172
QY	4139	TTCTTCAAGGGCATTCGACTTTTGGGACATTCGCGGACGAGAAGCGCCCTACATACCGGAA	4198
Db	4173	TTCTTCAAGGGCATTCGACTTTTGGGACATTCGCGGACGAGAAGCGCCCTACATACCGGAA	4232

[illegible]

Db 5313 GAGGAAACCGGAGGAAATTCACACTTATTCTCTGATAGCAACCGG 5360
|||||

RESULT 4

AC007821/c

LOCUS

DEFINITION

AC007821 AC007974

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

CELNIKER, S.E., ADAMS, M.D., KROMMILLER, B., TYLER, D., WAN, K.H.,
HOLT, R.A., EVANS, C.A., GOCAYNE, J.D., AMANATIDES, P.G., BRANDON, R.C.,
ROGERS, Y., AN, H., BALDWIN, D., BANZON, J., BEESON, K.Y., BUSAM, D.A.,
CARLSON, J.W., CENTER, A., CHAMPE, M., DAVENPORT, L.B., DIETZ, S.M.,
DODSON, K., DORSETT, V., DOUP, L.E., DOYLE, C., DRESNEK, D., FARFAN, D.,
FERRIERA, S., FRISSE, E., GALLIE, R.F., GARG, N.S., GEORGE, R.A.,
GONZALEZ, M., HOUCK, J., HOSKINS, R.A., HOSTIN, D., HOWLAND, T.J.,
IBEGWAM, C., JALALI, M., KRUSE, D., LI, P., MATTEI, B., MOSHREFI, A.,
MCINTOSH, T.C., MOY, M., MURPHY, B., NELSON, C., NELSON, K.A., NUNOO, J.,
PACLEB, J., PARAGAS, V., PARK, S., PATEL, S., PFEIFFER, B.,
PHOUANENONG, S., PITTMAN, G.S., PURI, V., RICHARDS, S., SCHEELER, F.,
STANLEY, M., STRONG, R., SVIRSKAS, R., TECTOR, C., WILLIAMS, S.M.,
ZAVARI, J.S., SMITH, H.O., RUBIN, G.M. and VENTER, J.C.
Sequencing of Drosophila chromosome 3R, region 99F-100A
Unpublished
2 (bases 1 to 163741)
CELNIKER, S.E., ABAZAYANI, A., ARCAINA, T.T., BAXTER, E., BLAZEJ, R.G.,
BUTENHOFF, C., CHAMPE, M., CHAVEZ, C., CHEW, M., CIESIOLKA, L.,
DOYLE, C.M., FARFAN, D.E., GALLIE, R., GEORGE, R.A., HARRIS, N.L.,
HOSKINS, R.A., HOUTSON, K.A., HUMMIST, S.R., KARRA, K., KEARNEY, L.,
KIM, E., LEE, B., LEWIS, S., LI, P., LOMOTAN, M.A., MAZDA, P.,
MOSHREFI, A.R., MOSHREFI, M., NIXON, K., PACLEB, J.M., PARK, S.,
PFEIFFER, B., POON, L., SEQUEIRA, A., SETHI, H., SNIR, E.,
SVIRSKAS, R.R., WAN, K.H., WEINBURG, T., ZHANG, R., ZIERAN, L.L. and
RUBIN, G.M.
Direct Submission
Submitted (14-JUN-1999) Drosophila Genome Center, Lawrence Berkeley
Laboratory, MS 64-121, Berkeley, CA 94720, USA
On or before Feb 24, 2001 this sequence version replaced
gi:5598720, gi:5670450.
Sequence submitted by:
Berkeley Drosophila Genome Project
Lawrence Berkeley National Laboratory, MS 64-121
Berkeley, CA 94720
This sequence was assembled using end sequences from a whole genome
shotgun and from subclones of this BAC and its neighboring clones.
For further information about this sequence, including its location
and relationship to other sequences, please visit our sequence
archive Web site (<http://www.fruitfly.org/sequence/>) or send email
to bdg@fruitfly.berkeley.edu.
Location/Qualifiers
1. 163741
/organism="Drosophila melanogaster"
/strain="y: cn bw sp"
/db_xref="taxon:7227"
/chromosome="3R"
/map="99F-100A"
/clone="BACR48102 (D803)"
/clone_lib="RPCI-98 (Roswell Park Cancer Institute
Drosophila melanogaster BAC library, partial EcoRI in
pBACE3.6)"
45739 a 36871 c 35638 g 45493 t

BASE COUNT

ORIGIN

Query Match

69.4%; Score 3969; DB 3; Length 163741;

Best Local Similarity 91.6%; Pred. No. 0;
Matches 4444; Conservative 22; Mismatches 17; Indels 368; Gaps 11;
QY 1223 CAGAAATTACACACCTCTGCGATACACGGGACCAACGAGGACGACGACTTACTCTCT 1282
|||||
DB 107198 CAGAAATTACACACCTCTGCGATACACGGGACCAACGAGGACGACGACTTACTCTCT 107139
|||||
QY 1283 GACTATACACACGCAAGCAGCGATGAGCGCGACCACTCCGCCCTCTCTCTGCTCCGGAC 1342
|||||
DB 107138 GACTATACACACGCAAGCAGCGATGAGCGCGACCACTCCGCCCTCTCTCTGCTCCGGAC 107079
|||||
QY 1343 GTGGTCATACCGCGCCGCCCAATTTAGTGTACGCCGCGGCTCCCATATATCCGTA 1402
|||||
DB 107078 GTGGTCATACCGCGCGCGCCCAATTTAGTGTACGCCGCGGCTCCCATATATCCGTA 107019
|||||
QY 1403 TCCGGTGTGGCGTGGAGTGTGGTGTGGCGAAGCAGGAGTGTSCCAAGATGATGACG 1462
|||||
DB 107018 TCCGGTGTGGCGTGGAGTGTGGTGTGGCGAAGCAGGAGTGTSCCAAGATGATGACG 106959
|||||
QY 1463 GGCCTAATGCCAAACAACTGTATCCGGAAGCGAGCATGAACGGGACACGCGGACAGT 1522
|||||
DB 106958 GGCCTAATGCCAAACAACTGTATCCGGAAGCGAGCATGAACGGGACACGCGGACAGT 106899
|||||
QY 1523 CACTACCTCGCTGCACTCCGGCTCTGAGTCTCGGAGCGCGTAGTCTCCGATCGGACAGC 1582
|||||
DB 106898 CACTACCTCGCTGCACTCCGGCTCTGAGTCTCGGAGCGCGTAGTCTCCGATCGGACAGC 106839
|||||
QY 1583 CCGCATTCGACACACACCCACGAGCGAGCTCGAGGACGGTGGTGAATCCAGTGAAT 1642
|||||
DB 106838 CCGCATTCGACACACACCCACGAGCGAGCTCGAGGACGGTGGTGAATCCAGTGAAT 106779
|||||
QY 1643 GTGGATTTTCTCCGTCGCAAGCGGTTTTCAGTGTAGTGTGCTCCACCGCGCGCGCCA 1702
|||||
DB 106778 GTGGATTTTCTCCGTCGCAAGCGGTTTTCAGTGTAGTGTGCTCCACCGCGCGCGCCA 106719
|||||
QY 1703 CGCAATTCACCGCTCTSCAGCGCGGCCACCGCCCCACCGCTAGTCCGCCACCGACAGC 1762
|||||
DB 106718 CGCAATTCACCGCTCTSCAGCGCGGCCACCGCCCCACCGCTAGTCCGCCACCGACAGC 106659
|||||
QY 1763 GGTAGTCTAAGCGGATCACCGCGCTTGAACACCGCGCGCGGATAGGCGGACCC 1822
|||||
DB 106658 GGTAGTCTAAGCGGATCACCGCGCTTGAACACCGCGCGGATAGGCGGACCC 106599
|||||
QY 1823 ACTCAGCGMGGCAACTCAGCTGTATTAACCCARAACCGGCTGAAGAACCCGCGACGACAG 1882
|||||
DB 106598 ACTCAGCGMGGCAACTCAGCTGTATTAACCCARAACCGGCTGAAGAACCCGCGACGACAG 106539
|||||
QY 1883 TTGACGACGACGTGAAGTCTCTGAACTTATACCCAGCGGAGGACGAGTGGAGTGTG 1942
|||||
DB 106538 TTGACGACGACGTGAAGTCTCTGAACTTATACCCAGCGGAGGACGAGTGGAGTGTG 106479
|||||
QY 1943 GAGCCACCGCGCTTACCTAATTCAGCGGAGCGGAGGACGACCGCGCGCGCA 2002
|||||
DB 106478 GAGCCACCGCGCTTACCTAATTCAGCGGAGCGGAGGACGACCGCGCGCGCA 106419
|||||
QY 2003 CCACCCAGTTACACGCGCTTCCATGAGTCCGCGGAGTCCGCCACACAAATCCCAACATCG 2062
|||||
DB 106418 CCACCCAGTTACACGCGCTTCCATGAGTCCGCGGAGTCCGCCACACAAATCCCAACATCG 106359
|||||
QY 2063 GACTACAGAAATCCCGAGCAGTGGGATATATCTCGGCCACCTCGCGCGGCTCGCGGAGC 2122
|||||
DB 106358 GACTACAGAAATCCCGAGCAGTGGGATATATCTCGGCCACCTCGCGCGGCTCGCGGAGC 106299
|||||
QY 2123 CCCTAATCTGTCTCTGCGCGCGCGCTGGCGGAGCGACCAACAGGAGTCTACACG 2182
|||||
DB 106298 CCCTAATCTGTCTCTGCGCGCGCGCTGGCGGAGCGACCAACAGGAGTCTACACG 106239
|||||
QY 2183 GCAGGAGTCACGCGGATCATCATGAGAGTGTGAAGCAGCAGGAGTGTCAAAAGGCC 2242
|||||
DB 106238 GCAGGAGTCACGCGGATCATCATGAGAGTGTGAAGCAGCAGGAGTGTCAAAAGGCC 106179
|||||
QY 2243 GTGCTGCAACAGCAGTGGCGCGCCCAATCGCCATCGAGTGTGCGGCCAGCAATTCACCA 2302
|||||

Db 106178 GTGCTGCAACAGCAGTGGGCGCCCAATCGCCATCGAGTGCCTCGGCCAGCAATTCACCA 106119
 Qy 2303 GTCCACGCTGCTGGCGGCTCCACCTCTTACCTCAGAACTCCGGGAGTGGTGCAGCAG 2362
 Db 106118 GTCCACGCTGCTGGCGGCTCCACCTCTTACCTCAGAACTCCGGGAGTGGTGCAGCAG 106059
 Qy 2363 CAGCAACAGGAGCAGGCGGCGCCACAGCAGAGCAGTACAGCAGCAGCAATCCAAACCA 2422
 Db 106058 CAGCAACAGGAGCAGGCGGCGCCACAGCAGCAGCAGTACAGCAGCAGCAATCCAAACCA 105999
 Qy 2423 CCAAGCGCMACCACACCGCCCTTGGTGGTCTGAAYAGCAAGCCCAATTCGCTGGAGCCA 2482
 Db 105998 CCAAGCGCCACCAACCGCCCTTGGTGGTCTGAATAGCAAGCCCAATTCGCTGGAGCCA 105939
 Qy 2483 CCGTCTATGCCAAGAGCAGTACAGGCCAAGCGGCCACGGTGGTARAGCAGCAGCAA --- 2539
 Db 105938 CCGTCTATGCCAAGAGCAGTACAGGCCAAGCGGCCACGGTGGTGCAGCAGCAGCAAACAG 105879
 Qy 2540 CAGCAGCAGCAGCAGCAGCAGTCCAGCAGCAGCAGGTGCAACAGCAGCAGCAGCAGCAG 2599
 Db 105878 CAGCAGCAGCAGCAGCAGCAGTCCAGCAGCAGCAGGTGCAACAGCAGCAGCAGCAGCAG 105819
 Qy 2600 CAACAGCAACTGCAGGCTTCCAGGCTGCTCCAGCAGCAGGCTCAGA --- 2645
 Db 105818 CAACAGCAACTGCAGGCTTCCAGGCTGCTCCAGCAGCAGGCTCAGAGGAGCGGGATCAA 105759
 Qy 2646 -GGGAGCGGATCAACGGGAGCGGAACGGGATCAGCAGAGCTGGCCAAACGGGAATCCT 2704
 Db 105758 CGGAGCGGATCAACGGGAGCGGAACGGGATCAGCAGAGCTGGCCAAACGGGAATCCT 105699
 Qy 2705 GGCGGAGAGTGTCTCGCGCGCCCTATCAGAGCAACAAACAAACAGCAGCAGCAGCAG 2764
 Db 105698 GGCGGAGAGTGTCTCGCGCGCCCTATCAGAGCAACAAACAAACAGCAGCAGCAGCAG 105639
 Qy 2765 AAACCGCGAGCTGCAACAAACAAACATACAGATAGCAGCAGCAACCTGGCGACGACA 2824
 Db 105638 AAACCGCGAGCTGCAACAAACAAACATACAGATAGCAGCAGCAACCTGGCGACGACA 105579
 Qy 2825 CCACCCATTCGCGCTGCAACAAACAAACATACAGATAGCAGCAGCAGCAGCAGCAGCAG 2884
 Db 105578 CCACCCATTCGCGCTGCAACAAACAAACATACAGATAGCAGCAGCAGCAGCAGCAGCAG 105519
 Qy 2885 GGCGGAGAGCAGTCCACCGGACACACCGCTCTCTGTCGACAGCAGTGTCAAGAAAGATC 2944
 Db 105518 GGCGGAGAGCAGTCCACCGGACACACCGCTCTCTGTCGACAGCAGTGTCAAGAAAGATC 105459
 Qy 2945 AAGCACGCTCGCCATCCCGGAGCGGAAGATCTCCAGAGGAAGAGGAGGAGCGC 3004
 Db 105458 AAGCACGCTCGCCATCCCGGAGCGGAAGATCTCCAGAGGAAGAGGAGGAGCGC 105399
 Qy 3005 AAGGAGTTCGCGATCAGGAGTACTCGCGCAAGCCTTCAAGTTCTTCATGGAGCAGCAGC 3064
 Db 105398 AAGGAGTTCGCGATCAGGAGTACTCGCGCAAGCCTTCAAGTTCTTCATGGAGCAGCAGC 105339
 Qy 3065 ATAGAGACGTGATCAGTCTGTCGAGCGGACGTATCCGAGAGATCAGCTGGAGAG 3124
 Db 105338 ATAGAGACGTGATCAGTCTGTCGAGCGGACGTATCCGAGAGATCAGCTGGAGAG 105279
 Qy 3125 GAGATGCAAAAGTGGGACTCCCGGATCAGACCCAAATCGAGATGAGGAAATCTGAAC 3184
 Db 105278 GAGATGCAAAAGTGGGACTCCCGGATCAGACCCAAATCGAGATGAGGAAATCTGAAC 105219
 Qy 3185 CAAAGGAGAGCAACTACATTCGATTTGAAGCGGCCCAAGATGGACAAGAGCATGTCGTC 3244
 Db 105218 CAAAGGAGAGCAACTACATTCGATTTGAAGCGGCCCAAGATGGACAAGAGCATGTCGTC 105159
 Qy 3245 AAACCTGAAGCCATTCGAGTGGGTGATTTGGCGAGGTAAAGCTGGTGAAGCAAAATCGAT 3304
 Db 105158 AAACCTGAAGCCATTCGAGTGGGTGATTTGGCGAGGTAAAGCTGGTGAAGCAAAATCGAT 105099
 Qy 3305 ACCTCGAAGCAATTTGATGCGATCAAAACCTTCGGGAAAGCGGAGCTTCTCAAGCGGAAT 3364
 Db 105098 ACCTCGAAGCAATTTGATGCGATCAAAACCTTCGGGAAAGCGGAGCTTCTCAAGCGGAAT 105039

Qy 3365 CAGGTGCGACACGTGAAGCGCGAGAGGATATCTCGGGAAGCCGACAAATAACTGGGTG 3424
 Db 105038 CAGGTGCGACACGTGAAGCGCGAGAGGATATCTCGGGAAGCCGACAAATAACTGGGTG 104979
 Qy 3425 GTGAAGTTGTACTACAGTCTCCAGGACAAAGATATCTGTACTTTGTGATGGACTACATA 3484
 Db 104978 GTGAAGTTGTACTACAGTCTCCAGGACAAAGATATCTGTACTTTGTGATGGACTACATA 104919
 Qy 3485 C----- 3485
 Db 104918 CCAGGTGAGTGGAAACTTAAGCAGTCCAAACGATGTGTATCGGAGTGCAAATGTCTATGTTG 104859
 Qy 3486 -----CAGTGGTGTATGTATGTCTGCTGCTCATCAAACTGGGCAATTTTCGAGGAG 3536
 Db 104858 GATTGTCGACAGGTGGTGTATGTATGTCTGCTGCTCATCAAACTGGGCAATTTTCGAGGAG 104799
 Qy 3537 AACTGGCCAGATTTCTACATCGCCGAGGTCACCTGGCCGCTGGAGCAGGCTTCACAAATGG 3596
 Db 104798 AACTGGCCAGATTTCTACATCGCCGAGGTCACCTGGCCGCTGGAGCAGGCTTCACAAATGG 104739
 Qy 3597 GCTTCATT----- 3604
 Db 104738 GCTTCATTACAGGTAATTAAGCAGTGTCTTCAAGTTATACAGTTCACATTTACATATGTT 104679
 Qy 3605 -----CACAGACATCAAGCTTGACAACTACTCATCGATAGGACCGGACACA 3653
 Db 104678 ATGGATTTCATCCACAGACATCAAGCTGACAACTACTCATCGATAGGACCGGACACA 104619
 Qy 3654 TAAAGCTCAGGACTTTGGCTGTGCAGGGATTTCCGATGGACGACCAACTCGAGTACT 3713
 Db 104618 TAAAGCTCAGGACTTTGGCTGTGCAGGGATTTCCGATGGACGACCAACTCGAGTACT 104559
 Qy 3714 ACCAGGAGAAC----- 3724
 Db 104558 ACCAGGAGACGGTGAAGTGAACCCAGCTAAGTGTGTTTAAACAGCTTTAATAATAT 104499
 Qy 3725 -----GGCAATCACTCGCCAGGACTTCGATGGAGCCCTGGGAGGAAATA 3768
 Db 104498 CATTCCTACTCTGATAGCAATCACTCGCCAGGACTTCGATGGAGCCCTGGGAGGAAATA 104439
 Qy 3769 CTCCGAGAACGGACCGAACCCACCGTGTCTGGAG----- 3802
 Db 104438 CTCCGAGAACGGACCGAACCCACCGTGTCTGGAGAGTGTGATTCGTACACTTAATTAAT 104379
 Qy 3803 -----AGCGACGGATCGCGGATCACCACAA 3827
 Db 104378 AGGATCGGCCCTTAATAAAGTATGCCAAGTTCGCGAGCGGATCGCGATCACCACAA 104319
 Qy 3828 GAGTCTGGCCCACTCGCTGGTGGGACCCCGAATCATATAGTCCCGAGGCTGCTGAGA 3887
 Db 104318 GAGTCTGGCCCACTCGCTGGTGGGACCCCGAATCATATAGTCCCGAGGCTGCTGAGA 104259
 Qy 3888 GAGTGGGTACACGACGCTGCGGACTTCTGGGCAACAGTCCGCTGGAAACGACAA 3947
 Db 104258 GAGTGGGTACACGACGCTGCGGACTTCTGGGCAACAGTCCGCTGGAAACGACAA 104199
 Qy 3948 TGGTGGGTACAGCCCTTCTGGGCAACAGTCCGCTGGAAACGACAA 3998
 Db 104198 TGGTGGGTACAGCCCTTCTGGGCAACAGTCCGCTGGAAACGACAA 104139
 Qy 3999 -----AGGTCA 4004
 Db 104138 CTTTGGATTCTTAAAGCAAGTGAATTAACAATGACTTCTTCAATTTACAGGTCA 104079
 Qy 4005 TCAACTGGGAAACCTGCATATTCGCGGAGCGGAGTATCCCGGAGGCTACGG 4064
 Db 104078 TCAACTGGGAAACCTGCATATTCGCGGAGCGGAGTATCCCGGAGGCTACGG 104019
 Qy 4065 ACTTGATAAGGAGGCTCTGTGCTGCGGTGCAACAGCGGCTGGGCAAGAGCTGGACGAG 4124
 Db 104018 ACTTGATAAGGAGGCTCTGTGCTGCGGTGCAACAGCGGCTGGGCAAGAGCTGGACGAG 103959

Li, J., Li, Z., Liang, Y., Lin, X., Liu, X., Mattei, B., McIntosh, T.C., McLeod, M.P., McPherson, D., Merkulov, G., Milshina, N.V., Mobarry, C., Morris, J., Moshrefi, A., Mount, S.M., Moy, M., Murphy, B., Murphy, L., Muzny, D.M., Nelson, D.L., Nelson, D.R., Nelson, K.A., Nixon, K., Nusskern, D.R., Pacle, J.M., Palazzolo, M., Pittman, C.S., Pan, S., Pollard, J., Puri, V., Reese, M.G., Reinert, K., Remington, K., Saunders, R.D., Scheeler, F., Shen, H., Shue, B.C., Siden-Kiamos, I., Simpson, M., Skupski, M.P., Smith, T., Spier, E., Spradling, A.C., Stapleton, M., Strong, R., Sun, E., Swirskas, R., Tector, C., Turner, R., Venter, E., Wang, A.H., Wang, X., Wang, Z.Y., Wasserman, D.A., Weinstein, G.M., Weissbach, J., Williams, S.M., Woodage, T., Worley, K.C., Wu, D., Yang, S., Yao, Q.A., Ye, J., Yeh, R.F., Zaveri, J.S., Zhan, M., Zhang, G., Zhao, Q., Zheng, L., Zheng, X.H., Zhong, F.N., Zhong, W., Zhou, X., Zhu, S., Zhu, X., Smith, H.O., Gibbs, R.A., Myers, E.W., Rubin, G.M., and Venter, J.C.
The genome sequence of *Drosophila melanogaster*.
Science 287 (5461), 2185-2195 (2000)
20196006
10731132
2 (bases 1 to 236109)
Adams, M.D., Celniker, S.E., Gibbs, R.A., Rubin, G.M. and Venter, C.J.
Direct Submission
Submitted (21-MAR-2000) Celera Genomics, 45 West Gude Drive, Rockville, MD, USA

FEATURES

Source

1..236109
Location/Qualifiers
/organism="Drosophila melanogaster"
/db_xref="taxon:7227"
/chromosome="3R"
/contig="CG1342"
/map="100A3-100A4"
/db_xref="FLYBASE:FBan0001342"
/db_xref="FLYBASE:FBgn0039795"
Join(20032..21133,21223..>2071)
/gene="CG1342"
/product="CT3018"
/db_xref="FLYBASE:FBan0001342"
/db_xref="FLYBASE:FBgn0039795"
Join(20033..21133,21223..>2071)
/gene="CG1342"
/note="CG1342 gene product"
/codon_start=1
/protein_id="AAF57075.1"
/db_xref="GI:7301970"
/db_xref="FLYBASE:FBan0001342"
/db_xref="FLYBASE:FBgn0039795"
/translation="MKAVLCATFTLAAIGQGLTQLEDENHSGFAGOVSLIALQL
LKFNKIDANQVHSPGLGASILAMLAASEGDTYSFEQVGVGPKDRTKLRLDAYKRIL
GSYQNRDAVALPSPOTWLYIYRNNSAREEFKLLKHVYVDVKDISREYDWNENPT
SLQSEKNDTSDTEGTSQNSKDVIGETLKRILKIDDEDIPAFSDGIGQEIIDE
ASKFDDVDKQVPRVPAQAEALLOKQKQATTESELESQPEETITLSVEKQEP
DMAAENPAQKQKSDQESQIKLNEETVOEEKLAKTMAAPALTAGEPKVL
PQKLENAVTAADGADEIMLALESHPFSVRNGARSLFQDDITLSANSITGR
SAGSKMLFLGLYRGWNPYQLRDGSDEFFMTNEDAVKAPMMHARGKFOVAD
LPQKARVLVETYSALICVLPDETEGLSDVI SOLQTSFLAKQKOFKELHIS
MPKFOVETSRSEAMLKQGLKKVSKTEAQLSLSEDDVHVEIVQVNRVNBEGG
SSANSLSAATMQARTPSVESTVLPVPEPELPVGFVRNPFATFIVDCQEQEVLA
SGLIYPERKEDLPVSIVIELEQS"
complement(<23504..>24751)
/gene="CG12069"
/map="100A3-100A4"
/db_xref="FLYBASE:FBan0012069"
/db_xref="FLYBASE:FBgn0039796"
complement(Join(<23504..23717,23771..24048,24117..24282,
24339..>24751))
/gene="CG12069"
/product="CT4640"
/db_xref="FLYBASE:FBan0012069"
/db_xref="FLYBASE:FBgn0039796"
complement(Join(23504..23717,23771..24048,24117..24282,
24339..24751))
/gene="CG12069"

gene

mRNA

CDS

gene

mRNA

CDS

/note="CG12069 gene product"
/codon_start=1
/protein_id="AAF57076.1"
/db_xref="GI:7301971"
/db_xref="FLYBASE:FBan0012069"
/db_xref="FLYBASE:FBgn0039796"
/translation="MGPPQQAQMHFSPKVDYIILLDKLRFDFNKKFATWTPSPSTGL
DDYELKATLGSGSGKQVLVRERESGGYIAQSKQSKDOIIVTKQVSHVSEKNNVLRSM
TFPNNVLIASYKDFDSLYLVPLIGGELFTYHRKYRKETEQAQYAAQVFLALEY
LHGSCLLYRDLKPFENIMMDKNGYLKVDGFAKKTERTMTLCGTPYLPPEIQQSP
YGTSDWAFAGVLVEFVAGHSPESAHNRDVMNSMYNKICADYKMPSYFSGARHLVD
HLLOVDLSKRFGLNLRNDRDKHEWEFKDVEWIPLLNQTYNAPVVPNISNPEDISNPD
KVSQDKPRKAKTMEFEAFADF"
complement(<24899..>26338)
/gene="Pka-C2"
/note="CG12066"
/map="100A3-100A4"
/db_xref="FLYBASE:FBan0012066"
/db_xref="FLYBASE:FBgn0000274"
complement(Join(<24899..25229,25297..25574,25629..25794,
25875..26104,26162..>26338))
/gene="Pka-C2"
/product="CT4596"
/note="Nucleotide sequence of the Celera sequence differs
from the published sequence for this transcript."
/db_xref="FLYBASE:FBan0012066"
/db_xref="FLYBASE:FBgn0000274"
complement(Join(24899..25229,25297..25574,25629..25794,
25875..26104,26162..26338))
/gene="Pka-C2"
/note="Pka-C2 gene product; Nucleotide sequence of the
Celera sequence differs from the published sequence for
this transcript"
/codon_start=1
/protein_id="AAF57077.1"
/db_xref="GI:7301972"
/db_xref="FLYBASE:FBan0012066"
/db_xref="FLYBASE:FBgn0000274"
/translation="NSQHTSQVFNSEKEDYNLLDNMSREFERNHQTSQPTNLEN
YITRAVLGNGSGTVMYLRKESGKNYIAAKMMSKEDLVRLKQVAVHNEKHVLAARF
PELIYLDSTKCFDYLILPLVNGELFYSYHRRVRKNEKHAHYAAQVLAALAHMH
KMLMYRDLKPNILLDQRYIKITDFGTRKVDGRTSLCGTPYLAPEIYQLRPYN
KSDWMAFGILYEFVAGSPFAIHNRILVMSKICIDYKMPSYFTSOLRSLVESL
MQVDTSKRLERRLRRESSVPGRLVHWSQPSGSHRPLPAHFRRRRSVRLRLV
QGSQVVPNKKPSRIVCLEPLNNVSEIVVFVSCVLSIRGAIIVIGCKSPFNTALLH
S"
/note="CG1340"
/map="100A5-100A5"
/db_xref="FLYBASE:FBan0001340"
/db_xref="FLYBASE:FBgn0039797"
Join(55458..56524,56585..>57119)
/gene="CG1340"
/product="CT3002"
/db_xref="FLYBASE:FBan0001340"
/db_xref="FLYBASE:FBgn0039797"
Join(55680..56524,56585..57119)
/gene="CG1340"
/note="CG1340 gene product"
/codon_start=1
/protein_id="AAF57078.1"
/db_xref="GI:7301973"
/db_xref="FLYBASE:FBan0001340"
/db_xref="FLYBASE:FBgn0039797"
/translation="MANRSQRDRFNRDNTYNNRNRDRHNRDRNRNRNFONY
RNRDRDRQRPVTEPFYIAYVGNLPGVLGVGMKIFSEFEVKNVRLIKDRETFE
KGYGVETETLAQLKALNCRKLDNFSAPLRIDIAHRRQNPAGPSVGGGAPAG
VCHERGCGVGGSGTGAAGNNPYORRNYRDRDSDSVGSHRRRPSNSNHQMSNS
TYNRSNNNSVGGGAGRGHPPORNGNGGGGPGFVNVNGGNTNFVONNRDR
RGHNPHSGSGSGYNNNGNTFGGNNHPVHGASASGCTGPGSSNQAVRDDDR
PKLVLPRTVTAPINSLAETKQALIFGKAKPRDSDSTPTVSSSPQDLIDASSGAPSDH
DHDFGLGN"

Fri Jan 17 11:17:08 2003

Db 122064 AAGGAGTTCGCGCATCAGGAGTACTCGCGCAGAGCCTTCAAGTCTTCATGAGCAGCAC 122005
 QY 3065 ATAGAGAACGTGATCAAGTCGTATCGCAGCGCAGCAGTATCGAAGAAATCAGCTGGAGAG 3124
 Db 122004 ATAGAGAACGTGATCAAGTCGTATCGCAGCGCAGCAGTATCGAAGAAATCAGCTGGAGAG 121945
 QY 3125 GAGATGCACAAAGTGGGACTGCCCGATCAGACCCCAATCGAGATGAGAAATGCTGAAC 3184
 Db 121944 GAGATGCACAAAGTGGGACTGCCCGATCAGACCCCAATCGAGATGAGAAATGCTGAAC 121985
 QY 3185 CAAAGGAGACACTACATTCGATCAAGCGCGCCAGATGACAGAGCATGTTCTGTC 3244
 Db 121884 CAAAGGAGACACTACATTCGATCAAGCGCGCCAGATGACAGAGCATGTTCTGTC 121825
 QY 3245 AAAGTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAAAGCTGGTGAACAAATCGAT 3304
 Db 121824 AAAGTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAAAGCTGGTGAACAAATCGAT 121765
 QY 3305 ACCTCGAACCATTTGATGCGATGAACACCTCGGAAAGCGGACGTTCTCAAGCGGAAT 3364
 Db 121764 ACCTCGAACCATTTGATGCGATGAACACCTCGGAAAGCGGACGTTCTCAAGCGGAAT 121705
 QY 3365 CAGGTGGCACACCTGAAGCGCGGAGGATATCCTCGCGGAAGCGCAATAAAGTGGTG 3424
 Db 121704 CAGGTGGCACACCTGAAGCGCGGAGGATATCCTCGCGGAAGCGCAATAAAGTGGTG 121645
 QY 3425 GTGAAGTGTACTACAGCTTCAGGACAGGATATCTGTACTTGTGATGGACTACATA 3484
 Db 121644 GTGAAGTGTACTACAGCTTCAGGACAGGATATCTGTACTTGTGATGGACTACATA 121585
 QY 3485 C----- 3485
 Db 121584 CCAAGTGGAGTGAACCTTAAGCAGTCCAAGATGTATCGGAGTGAACATGTCAATGTG 121525
 QY 3486 -----CAGGTGGTGTATCGATCTGCTCATCAAACTGGGCAATTTTCAGAGAGG 3536
 Db 121524 GATTTGTCAGACAGTGTGTATCGTCTCATCAAACTGGGCAATTTTCAGAGAGG 121465
 QY 3537 AACTGGCCACATTTACATCGCCGAGGTCACTTCGCCCGTGGACAGCGTTTCAAAATGG 3596
 Db 121464 AACTGGCCACATTTACATCGCCGAGGTCACTTCGCCCGTGGACAGCGTTTCAAAATGG 121405
 QY 3597 GCTTCATT----- 3604
 Db 121404 GCTTCATTCAAGGTAATTAAGCAGTTCATGCTTATCAGCTACATTTACATATTTGT 121345
 QY 3605 -----CACAGAGACATCAAGCTCGACACATCTCATCTAGGAGGACACA 3653
 Db 121344 ATGGATTCATCCACAGACATCAAGCTCGACACATCTCATCTAGGAGGACACA 121285
 QY 3654 TAAAGCTCACCGACTTTGGCTGTGACGGGATTCGGATGACGACCACTCGAAGTACT 3713
 Db 121284 TAAAGCTCACCGACTTTGGCTGTGACGGGATTCGGATGACGACCACTCGAAGTACT 121225
 QY 3714 ACCAGGAGAAC----- 3724
 Db 121224 ACCAGGAGAACGGTGAGATGGAACCCAGCTAACGTAGTGTAAACAGCTTTAATAATAT 121165
 QY 3725 -----GGCATCTCCGCGCAGGCTCGATGAGCGCTTGGAGGAATA 3768
 Db 121164 CATTCCTACTCTGATAGGCAATCACTCGCGCAGGACTCGATGAGCGCTTGGAGGAATA 121105
 QY 3769 CTCGAGACGACGACGAGCCACCGCTGCTGAGAGGTGAGTTCGTACACTTAATTAAT 121045
 Db 121104 CTCGAGACGACGACGAGCCACCGCTGCTGAGAGGTGAGTTCGTACACTTAATTAAT 3827
 QY 3803 -----AGCGGAGGTGCGGATCAGCAAA 3877
 Db 121044 AGGATCGCGGTTTAAATAAAGTATCGGAACTCGCAGCGGAGGTGCGGATCAGCAAA 120985
 QY 3828 GAGTCTGGCCCACTCGCTGGTGGGACCCCGAAGTATACATAGCTCCCGAGGTGCTGGAGA 3887
 Db 120984 GAGTCTGGCCCACTCGCTGGTGGGACCCCGAAGTATACATAGCTCCCGAGGTGCTGGAGA 120925

QY 3888 GRAGTGGGTACACGACGCTGTGCGACTACTGAGCGTGGCGTTCATCTTATGAGATGC 3947
 Db 120924 GAAGTGGGTACACGACGCTGTGCGACTACTGAGCGTGGCGTTCATCTTATGAGATGC 120865
 QY 3948 TGGTGGGTACGCGCCCTTTCTGCGCAACAGTCCGCTGGAAACGCAACAAA----- 3998
 Db 120864 TGGTGGGTACGCGCCCTTTCTGCGCAACAGTCCGCTGGAAACGCAACAAAAGTTCGCT 120805
 QY 3999 -----AGGTCA 4004
 Db 120804 CTTTGGATTCCTTATAAAGCAAGTGCATTTACAAATGACTTGTCTTACAGGTCA 120745
 QY 4005 TCAACTGGGAGAAACSCGTGCATATTCGCGCGCAGCGCGAGTTATCCCGGAGGCTACGG 4064
 Db 120744 TCAACTGGGAGAAACSCGTGCATATTCGCGCGCAGCGCGAGTTATCCCGGAGGCTACGG 120685
 QY 4065 ACTTGAATAAGGAGGCTGTGCGCGCTGACAAAGCGGCTGGCAAGAGCGTGACAGAGG 4124
 Db 120684 ACTTGAATAAGGAGGCTGTGCGCGCTGACAAAGCGGCTGGCAAGAGCGTGACAGAGG 120625
 QY 4125 TCAAGAGCCACGACTTCTTCAAGGSCATCGACTTTTGGGACATCGGAAAGCAGAAAGCGC 4184
 Db 120624 TCAAGAGCCACGACTTCTTCAAGGSCATCGACTTTTGGGACATCGGAAAGCAGAAAGCGC 120565
 QY 4185 CCTTACATACCGGAAATCAAGCACCACCGGACATCAACCTTTTCATCCCGTGGATCGG 4244
 Db 120564 CCTTATATCCGGAATCAAGCACCACCGGACATCAACCTTTTCATCCCGTGGATCGG 120505
 QY 4245 AGAAGCTGGCTGATGATCCACCATGAGCAGCGGCGATGATGTCGACCAAGATGACC 4304
 Db 120504 AGAAGCTGGCTGATGATCCACCATGAGCAGCGGCGATGATGTCGACCAAGATGACC 120445
 QY 4305 GCACCTTCCACGCGCTTTTCGAATTTTACCTTCGCTGCTTTCGACGACGAGCAGCGC 4364
 Db 120444 GCACCTTCCACGCGCTTTTCGAATTTTACCTTCGCTGCTTTCGACGACGAGCAGCGC 120385
 QY 4365 CGGATATGAGCGAGTACAGCGCGCTTTTACGCTGAAATGATGCTCTCCATGTGCC 4424
 Db 120384 CGGATATGAGCGAGTACAGCGCGCTTTTACGCTGAAATGATGCTCTCCATGTGCC 120325
 QY 4425 AACACCAACACCCCGCCCGGAAATCATTTGTAGTCAAAATAGTCAAAAAGGGGATAGAA 4484
 Db 120324 AACACCAACACCCCGCCCGGAAATCATTTGTAGTCAAAATAGTCAAAAAGGGGATAGAA 120265
 QY 4485 ACCATGAGTGGGCTTCGATTTAAAGGAAAGCGTGGCTATAGAAATGAAACTATCTATATA 4544
 Db 120264 ACCATGAGTGGGCTTCGATTTAAAGGAAAGCGTGGCTATAGAAATGAAACTATCTATATA 120205
 QY 4545 CATTATATAAATTTAGAGACAGTAGAGCGGAGCTAGTATATACATAAATAAATA 4604
 Db 120204 CATTATATAAATTTAGAGACAGTAGAGCGGAGCTAGTATATACATAAATAAATA 120145
 QY 4605 TACATATATTTGAT 4663
 Db 120144 TACATATATTTGAT 120085
 QY 4664 AAAACGAGCGCGAGTAGAGATGAAA- CGAGAGGAGCGAGTCCAGACCTTCGACCTTTAAC 4722
 Db 120084 AAAACGAGCGCGAGTAGAGATGAAAACCGAGAGGAGTCCAGACCTTCGACCTTTAAC 120025
 QY 4723 TGAACATAGTATATCTTGTGCTACTTCCACACCAATAATATATATATATATATATATAT 4782
 Db 120024 TGAACATAGTATATCTTGTGCTACTTCCACACCAATAATATATATATATATATAT 119965
 QY 4783 AGAATTCAAAAGGKCAACTGGAATTCGAATTCGAGTCTGAGTCTGAGTCTGAGTCTGAGT 4842
 Db 119964 AGAATTCAAAAGGKCAACTGGAATTCGAATTCGAGTCTGAGTCTGAGTCTGAGTCTGAGT 119915
 QY 4843 AAAGCAAAACAAACGCTTTAACTAAATGAGCAGCGAATTTTACCAACCACTTCACTCC 4902
 Db 119914 AAAGCAAAACAAACGCTTTAACTAAATGAGCAGCGAATTTTACCAACCACTTCACTCC 119855

QY 4903 TCTCTTTTCCACCTCCGATCGTGGCCGGAATTCGAACCTCAGAGCTGGTTGCATCG 4962
 Db 119854 TCTCTTTTCCACCTCCGATCGTGGCCGGAATTCGAACCTCAGAGCTGGTTGCATCG 119795
 QY 4963 GCCATCCCATTTACCTCCATTCAGATTCGAGGAGTGGAGTGGAGAACGAAC 5022
 Db 119794 GCCATCCCATTTACCTCCATTCAGATTCGAGGAGTGGAGTGGAGAACGAAC 119735
 QY 5023 GGAGACCAAAAGTCGACGCGGAGGATTAAGCGGGTCTTAAGGCTTAATCTAAATCTA 5082
 Db 119734 GGAGACCAAAAGTCGACGCGGAGGATTAAGCGGGTCTTAAGGCTTAATCTAAATCTA 119675
 QY 5083 AACTGGGAGAACAGGACCCATCTA-----TGTCTCTGATTCACCAATTCGCTATC 5131
 Db 119674 AACTGGGAGAACAGGACCCATCTA-----TGTCTCTGATTCACCAATTCGCTATC 119615
 QY 5132 ACTGCTCTTCANCTGWTAGGACCCACCCGCCCTC-----CCCATCCAAAGAACAA 5187
 Db 119614 ACTGCTCTTCANCTGWTAGGACCCACCCGCCCTC-----CCCATCCAAAGAACAA 119555
 QY 5188 ACTTAGACCTAGCTATGTGAAGAGCTAGCAATGTTAGACCAACTGTTGTAATGCCAAAT 5247
 Db 119554 ACTTAGACCTAGCTATGTGAAGAGCTAGCAATGTTAGACCAACTGTTGTAATGCCAAAT 119495
 QY 5248 GAAATGTTTACGCCCCACGAGGAAACGCGGGGAAATTTCAACACTTATCTCTGATAGC 5307
 Db 119494 GAAATGTTTACGCCCCACGAGGAAACGCGGGGAAATTTCAACACTTATCTCTGATAGC 119435
 QY 5308 AAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5367
 Db 119434 AAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 119380
 QY 5368 TTAATGTAATATGTAAGAACACGTTAATGTAATGTAATGTAATGTAATGTAATGTAAT 5427
 Db 119379 TTAATGTAATATGTAAGAACACGTTAATGTAATGTAATGTAATGTAATGTAATGTAAT 119320
 QY 5428 TAAGATGTTTATGTTATAGACCGCTAACCGTATCTAGTTTAACTTCTTAACACTAAG 5487
 Db 119319 TAAGATGTTTATGTTATAGACCGCTAACCGTATCTAGTTTAACTTCTTAACACTAAG 119260
 QY 5488 CGAGAGTACAGTACATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5547
 Db 119259 CGAGAGTACAGTACATGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 119200
 QY 5548 GAAACGATTTGTTTCTCTTTAATAGCTTCAGTTGTTGTTGTTGTTGTTGTTGTTGTT 5607
 Db 119199 GAAACGATTTGTTTCTCTTTAATAGCTTCAGTTGTTGTTGTTGTTGTTGTTGTTGTT 119140
 QY 5608 GACTTATATATAGTCCATCTGAATATTCGTGGATGAGGCTATTTTAAATGTCAGATCGA 5667
 Db 119139 GACTTATATATAGTCCATCTGAATATTCGTGGATGAGGCTATTTTAAATGTCAGATCGA 119080
 QY 5668 GCTAATGGAAGAAATACAAACAACTCTGTGCTGGCTGGCCCAATTAGTTTAC 5720
 Db 119079 GCTAATGGAAGAAATACAAACAACTCTGTGCTGGCTGGCCCAATTAGTTTAC 119027

RESULT 6
 AC012975/c
 LOCUS AC012975 28133-bp DNA linear HTG 03-NOV-1999
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS *** in ordered
 pieces
 ACCESSION AC012975
 VERSION AC012975.1 GI:6223348
 KEYWORDS HTG; HFGS_PHASE2
 SOURCE Drosophila melanogaster
 ORGANISM Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila
 1 (bases 1 to 28133)
 AUTHORS Adams,M. and Venter,J.C.
 TITLE Direct Submission

JOURNAL Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 COMMENT This sequence was identified as CDM:10211994 by the submitter.
 For further information on this sequence you may e-mail to:
 fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.
 FEATURES
 source location/Qualifiers
 1..28133
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 BASE COUNT 7406 a 6301 c 6224 g 8202 t
 ORIGIN
 Query Match 51.8%; Score 2965.2; DB 2; Length 28133;
 Best Local Similarity 90.3%; Pred. No. 0;
 Matches 3369; Conservative 16; Mismatches 9; Indels 338; Gaps 7;
 QY 1223 CAGATTATCACACCTCTGCGATACAGCGGACGACGAGCGAAGCATGCTACTCT 1282
 Db 3732 CAGATTATCACACCTCTGCGATACAGCGGACGACGAGCGAAGCATGCTACTCT 3673
 QY 1283 GACTATCACACGCGCAAGCAGCGGATGAGCGCGCCACCTCCGCCCTCTCTGCTCGGAC 1342
 Db 3672 GACTATCACACGCGCAAGCAGCGGATGAGCGCGCCACCTCCGCCCTCTCTGCTCGGAC 3613
 QY 1343 GTGGTCATACCGCGCGCGCGCCCATTTAGTGTACGCGCGGAGCGCGGTCTCCATATCCGTA 1402
 Db 3612 GTGGTCATACCGCGCGCGCGCCCATTTAGTGTACGCGCGGAGCGCGGTCTCCATATCCGTA 3553
 QY 1403 TCCGCTGTGGCGGTGGAGTGTGGTGTGGCGAGCGAGCTGTSCCAAGATGATGAGC 1462
 Db 3552 TCCGCTGTGGCGGTGGAGTGTGGTGTGGCGAGCGAGCTGTSCCAAGATGATGAGC 3493
 QY 1463 GCCTTAATCCCAACAACTGATTCGGAAGCGAGGATCGAAGCGGACACGCGGAGCAGT 1522
 Db 3492 GCCTTAATCCCAACAACTGATTCGGAAGCGAGGATCGAAGCGGACACGCGGAGCAGT 3433
 QY 1523 CACTACCTGCGTGCAGTCCGGCTCTGGACTCCGAGCGCGGTAGTCCCGATCGGACAGC 1582
 Db 3432 CACTACCTGCGTGCAGTCCGGCTCTGGACTCCGAGCGCGGTAGTCCCGATCGGACAGC 3373
 QY 1583 CCGCATTCGACACACACCCACCGAGCTCGAGGACGCGTGGTAAATCCAGGTGGAAT 1642
 Db 3372 CCGCATTCGACACACACCCACCGAGCTCGAGGACGCGTGGTAAATCCAGGTGGAAT 3313
 QY 1643 GGTGATTTTCCTCGTTCGCAAGCGGTTTCAGTGTAGTGTGCTCCACCGCGCGCGCCA 1702
 Db 3312 GGTGATTTTCCTCGTTCGCAAGCGGTTTCAGTGTAGTGTGCTCCACCGCGCGCGCCA 3253
 QY 1703 CGCAATCCACCGCTTSCAGCGCGCGCACCGCCCCACCGCCAGTGCAGCGCCACAGCCAG 1762
 Db 3252 CGCAATCCACCGCTTSCAGCGCGCGCACCGCCCCACCGCCAGTGCAGCGCCACAGCCAG 3193
 QY 1763 GGTGATTCGAGCGGATCACCGCGCTGAAACACCGCGCGCGCGCGGATGCGCCACCC 1822
 Db 3192 GGTGATTCGAGCGGATCACCGCGCTGAAACACCGCGCGCGCGGATGCGCCACCC 3133
 QY 1823 ACTCAGCGGCGCAACTCACCTGTAATACCCAAACGGGCTGAAGAACCCGCGAGCAGAG 1882
 Db 3132 ACTCAGCGGCGCAACTCACCTGTAATACCCAAACGGGCTGAAGAACCCGCGAGCAGAG 3073
 QY 1883 TTGACGCGAGCTGAAGTCCCTGAACCTTATACCGCGGAGCGAGTGGAGCTGCTG 1942
 Db 3072 TTGACGCGAGCTGAAGTCCCTGAACCTTATACCGCGGAGCGAGTGGAGCTGCTG 3013
 QY 1943 GAGCCACCGCGCGCTACCTAATTCAGCGGAGCGCGGAGGAGCAGCCGCGCGCGCA 2002
 Db 3012 GAGCCACCGCGCGCTACCTAATTCAGCGGAGCGCGGAGGAGCAGCCGCGCGCGCA 2953
 QY 2003 CCACCCAGTTACAGCGGCTTCCATGTCAGTGCAGTGCAGCCACACAAATCCCAACATCG 2062

Db 2952 CCACCCAGTTACAGGCTTCCATGAGTCGCGGAGTCCGCCACACAAATCCACAAATCG 2893
 QY 2063 GACTACAGAAATCCCGAGCAGTGGGATATACCTCGGCCACCTCGCGGCTCGCGAGC 2122
 Db 2892 GACTACAGAAATCCCGAGCAGTGGGATATACCTCGGCCACCTCGCGGCTCGCGAGC 2833
 QY 2123 CCATTAAGTGTGCTGCTGCGGCGGCGGCTGCGGAGGAGCAACACAGAGTCTACCG 2182
 Db 2832 CCATTAAGTGTGCTGCTGCGGCGGCGGCTGCGGAGGAGCAACACAGAGTCTACCG 2773
 QY 2183 GCCAGGAGTACAGCCGATCATCATGAGTGTGAAGAGCAGCAGGTCCAAAAGCC 2242
 Db 2772 GCCAGGAGTACAGCCGATCATCATGAGTGTGAAGAGCAGCAGGTCCAAAAGCC 2713
 QY 2243 GTGCTGCAAAACAGCAGTGGCGGCGGCAATGCCATCGAGTGCCTCGGCCACCAATTCACCA 2302
 Db 2712 GTGCTGCAAAACAGCAGTGGCGGCGGCAATGCCATCGAGTGCCTCGGCCACCAATTCACCA 2653
 QY 2303 GTCCAGTGTGCGGCTCCACCTCTTACCCCTCAGAAGTCCGGGAGTGGTGCAGCAG 2362
 Db 2652 GTCCAGTGTGCGGCTCCACCTCTTACCCCTCAGAAGTCCGGGAGTGGTGCAGCAG 2593
 QY 2363 CAGCAACAGCAGCAGCGGCGGCGGCAATGCCATCGAGTGCCTCGGCCACCAATTCACCA 2422
 Db 2592 CAGCAACAGCAGCAGCGGCGGCGGCAATGCCATCGAGTGCCTCGGCCACCAATTCACCA 2533
 QY 2423 CCAACGCCMACCAACCGGCTTGTGGTGTGAAGCAAGCCCAATTCCTCGGAGCA 2482
 Db 2532 CCAACGCCMACCAACCGGCTTGTGGTGTGAAGCAAGCCCAATTCCTCGGAGCA 2473
 QY 2483 CCGTCTATGCCAAGAGCAGTGCAGGCGGCGGCAATTCCTCGGAGCAAGCAAGCA 2539
 Db 2472 CCGTCTATGCCAAGAGCAGTGCAGGCGGCGGCAATTCCTCGGAGCAAGCAAGCA 2413
 QY 2540 CAGCAGCAGCAGCAGGCTGCAGGAGCAGGAGTGCAGCAGCAGCAGCAGCAGCAG 2599
 Db 2412 CAGCAGCAGCAGCAGGCTGCAGGAGCAGGAGTGCAGCAGCAGCAGCAGCAGCAG 2533
 QY 2600 CAACAGCACTCAGGCTTGGGCTGCTCAGGAGCAGGCTCAGA 2645
 Db 2352 CAACAGCACTCAGGCTTGGGCTGCTCAGGAGCAGGCTCAGA 2293
 QY 2646 -GGGAGCGGATCAACGGGAGCGGAGCGGATCAGCAGAGAGTGGCCACGGAAATCCT 2704
 Db 2292 CGGAGCGGATCAACGGGAGCGGAGCGGATCAGCAGAGAGTGGCCACGGAAATCCT 2233
 QY 2705 GCGGCGAGATGCTTCCGCGCGGCGGCTATCAGAGCAACAAACACAGCAGATC 2764
 Db 2232 GCGGCGAGATGCTTCCGCGCGGCGGCTATCAGAGCAACAAACACAGCAGATC 2173
 QY 2765 AAACCGCGAGCTGCAACAAACAAACATACAGATAGCAACAGCACTGGCGAGCA 2824
 Db 2172 AAACCGCGAGCTGCAACAAACAAACATACAGATAGCAACAGCACTGGCGAGCA 2113
 QY 2825 CCACCCATTCGCTGCCAATAACAACTCTTCCAAACCGGCGGGAATAGCTCG 2884
 Db 2112 CCACCCATTCGCTGCCAATAACAACTCTTCCAAACCGGCGGGAATAGCTCG 2053
 QY 2885 GCGGCGAGCAAGGATCCACCGGCGGCGGCTTCCAAACCGGCGGGAATAGCTCG 2944
 Db 2052 GCGGCGAGCAAGGATCCACCGGCGGCGGCTTCCAAACCGGCGGGAATAGCTCG 1993
 QY 2945 AAGCAGCCTCGCCATTCGCGGAGCGCAAGAGATCTCCAAAGAGAGGAGGAGCGC 3004
 Db 1992 AAGCAGCCTCGCCATTCGCGGAGCGCAAGAGATCTCCAAAGAGAGGAGGAGCGC 1933
 QY 3005 AAGAGTTCGCGATCAGGAGTACTCGCGCAGCGCTTCAAGTCTTATGAGCAGCAGC 3064
 Db 1932 AAGAGTTCGCGATCAGGAGTACTCGCGCAGCGCTTCAAGTCTTATGAGCAGCAGC 1873
 QY 3065 ATAGAAACGTGATCAAGTCTGATCCAGCAGCAGTATCGAAGANTCAGTGGAGAG 3124

Db 1872 ATAGAGAACGTGATCAAGTCTGATCCAGCGCACGTATCCAGAGATCAGCTGGAGAG 1813
 QY 3125 GAGATGCACAAAGTGGAGTCCCGATCAGACCCAAATCGAGATGAGAAATGCTGAAC 3184
 Db 1812 GAGATGCACAAAGTGGAGTCCCGATCAGACCCAAATCGAGATGAGAAATGCTGAAC 1753
 QY 3185 CAAAAGGAGAGCAACTACATTCGATTGAAGCGCGCAAGATGGCAAGAGCATCTCGTC 3244
 Db 1752 CAAAAGGAGAGCAACTACATTCGATTGAAGCGCGCAAGATGGCAAGAGCATCTCGTC 1693
 QY 3245 AAATGAGAGCCATTCGAGTGGTGCATTCGGAGGTAAAGTGGTGAAGCAAAATCGAT 3304
 Db 1692 AAATGAGAGCCATTCGAGTGGTGCATTCGGAGGTAAAGTGGTGAAGCAAAATCGAT 1633
 QY 3305 ACCTCGAAACCATTTGATGCCATGAAACCCCTCGGAAAGCGGAGCTTCTCAAGCGGAAT 3364
 Db 1632 ACCTCGAAACCATTTGATGCCATGAAACCCCTCGGAAAGCGGAGCTTCTCAAGCGGAAT 1573
 QY 3365 CAGTGGCAGACGTGAAGCGGAGAGGAGTATCTCGGAAAGCGGAGCAATAAATGGGTG 3424
 Db 1572 CAGTGGCAGACGTGAAGCGGAGAGGAGTATCTCGGAAAGCGGAGCAATAAATGGGTG 1513
 QY 3425 GTGAAGTGTACTACAGCTTCCAGGACAAAGATAAATCTGATCTTGTGATGGATACATA 3484
 Db 1512 GTGAAGTGTACTACAGCTTCCAGGACAAAGATAAATCTGATCTTGTGATGGATACATA 1453
 QY 3485 C----- 3485
 Db 1452 CAGTGGTGAAGCAATTAAGCAGTCCACGAGTGTCTCGGAGTCCAAATGTCATTTGT 1393
 QY 3486 -----CAGTGGTGAAGCAATTAAGCAGTGTCTCGGAGTCCAAATGTCATTTGT 3536
 Db 1392 GATTGTCAGCAGGTGGTGTCTGATGCTGCTCATCAAACTGGGCAATTTTCGAGGAG 1333
 QY 3537 AACTGCCAGATTTACATCGCGGAGTCACTGCGCGTGGACAGCGTTCACAAATGG 3596
 Db 1332 AACTGCCAGATTTACATCGCGGAGTCACTGCGCGTGGACAGCGTTCACAAATGG 1273
 QY 3597 GCTTCAT----- 3604
 Db 1272 GCTTCATCAGGTAATTAAGCAATTAAGCAATTAAGCAATTAAGCAATTAATTTGT 1213
 QY 3605 -----CACAGACATCAAGCCCTGCAACATCTCATCTAGGAGCGGAGCAACA 3653
 Db 1212 ATGGATTATCAGCAGAGATCAAGCCCTGCAACATCTCATCTAGGAGCGGAGCAACA 1153
 QY 3654 TAAAGTCAACCACTTTGCGCTGTCAGCGGATTCGATGGAGCAGCAACTCGAAGTACT 3713
 Db 1152 TAAAGTCAACCACTTTGCGCTGTCAGCGGATTCGATGGAGCAGCAACTCGAAGTACT 1093
 QY 3714 ACCAGGAGAAC----- 3724
 Db 1092 ACCAGGAGAACCGTGCAGATGGNACCCAGCTAACGTAGTGTAAACAGCTTTAATAATAT 1033
 QY 3725 -----GGCAATCACTCCCGCAGGAGTTCGATGGAGCGCTGGGAGGAATA 3768
 Db 1032 CATTCCTACTGTAGGCAATCACTCGCGCAGGAGTTCGATGGAGCGCTGGGAGGAATA 973
 QY 3769 CTCCGAGAACCGGAGCGGAGCGGAGTTCGAG----- 3802
 Db 972 CTCCGAGAACCGGAGCGGAGCGGAGTTCGAG----- 913
 QY 3803 -----AGGCGAGGATCGCGATCACCATA 3827
 Db 912 AGGATCGCGCTTAATAAAGTATGCGGAATCGCAGCGGAGTTCGCGATCACCATA 853
 QY 3828 GAGTCTCGGCGCTTCCGCTGGTGGGACCCCGCAACTACATAGTCCCGAGGAGTCTGGAGA 3887
 Db 852 GAGTCTCGGCGCTTCCGCTGGTGGGACCCCGCAACTACATAGTCCCGAGGAGTCTGGAGA 793
 QY 3888 GAGTGGGTACACAGCAGCTGTGGAGTACTTGGAGCGTGGCGCTCATCTTATGAGATGC 3947
 Db 792 GAGTGGGTACACAGCAGCTGTGGAGTACTTGGAGCGTGGCGCTCATCTCTATGAGATGC 733

QY 3948 TGTGGTGCAGCGCCCTTTCTGGCCACAGTCCGCTGGAACGCAACAA----- 3998
 Db 732 TGTGGTGCAGCGCCCTTTCTGGCCACAGTCCGCTGGAACGCAACAAAGGTTGCT 673
 QY 3999 -----AGTCA 4004
 Db 672 CTTTGATTCCTTATAAGCAAGTCAATTACAACAATGACTTGTCTTATACAGTCA 613
 QY 4005 TCAACTGGGAGAAACSCGTGATATTCGCGCAGCGCCGAGTTATCCGCGAGGCTACG 4064
 Db 612 TCNACTGGGAGAAACCCCTGATATTCGCGCAGCGCCGAGTTATCCGCGAGGCTACG 553
 QY 4065 ACTTGATAAGGAGGCTCTGCTGGCTGGCTGACAAGCGGCTGGGCAAGAGCGTGACGAG 4124
 Db 552 ACTTGATAAGGAGGCTCTGCTGGCTGGCTGACAAGCGGCTGGGCAAGAGCGTGACGAG 493
 QY 4125 TCAAGAGCCAGGACTTCTTCAAGGGCATCGACTTTCCGGACATGCGGAAAGCAGAGCGC 4184
 Db 492 TCAAGAGCCAGGACTTCTTCAAGGGCATCGACTTTCCGGACATGCGGAAAGCAGAGCGC 433
 QY 4185 CTTACATACCGGAAATCAAGCAGCCACGACATCCCACTTTGATCCGCTGGATCCG 4244
 Db 432 CTTATATACCGGAAATCAAGCAGCCACGACATCCCACTTTGATCCGCTGGATCCG 373
 QY 4245 AGAAGCTGGCTCGAATGACTCCACCATGAGCAGCGCGGATGATGTCGACCAAGATGACC 4304
 Db 372 AGAAGCTGGCTCGAATGACTCCACCATGAGCAGCGCGGATGATGTCGACCAAGATGACC 313
 QY 4305 GCACATTCACGCGCTTTTCGAATTTACCTTCGCTGCTTTCGACGACAGCAGCGCG 4364
 Db 312 GCACATTCACGCGCTTTTCGAATTTACCTTCGCTGCTTTCGACGACAGCAGCGCG 253
 QY 4365 CGGATATGAGGAGCATGAGCGCGGTTTACGCTGGAATGATGATGCTCTCCATGTCGC 4424
 Db 252 CGGATATGAGGAGCATGAGCGCGGTTTACGCTGGAATGATGATGCTCTCCATGTCGC 193
 QY 4425 AACACCAACACCCCGCCCGAATCAATTTAGTCAAAATAGTACAAAAAGGGATAGAA 4484
 Db 192 AACACCAACACCCCGCCCGAATCAATTTAGTCAAAATAGTACAAAAAGGGATAGAA 133
 QY 4485 ACCATTGAGTGGCTTGATTTGAAAGGAGCGTGTGCTATAGAACTATCTATATA 4544
 Db 132 ACCATTGAGTGGCTTGATTTGAAAGGAGCGTGTGCTATAGAACTATCTATATA 73
 QY 4545 CATTATATAAATTATAGGAGACAGTAGAGCGGAGCTACGTATATACATAAATAATA 4604
 Db 72 CATTATATAAATTATAGGAGACAGTAGAGCGGAGCTACGTATATACATAAATAATA 13
 QY 4605 TACATATATTG 4616
 Db 12 TACATATATTG 1
 RESULT 7
 AC013000/c
 LOCUS
 DEFINITION Drosophila melanogaster, *** SEQUENCING IN PROGRESS ***
 ACCESSION AC013000
 VERSION AC013000.1 GI:6223329
 KEYWORDS HTG; HTGS_PHASE2.
 SOURCE Drosophila melanogaster.
 ORGANISM Drosophila melanogaster
 Eukaryota; Metazoa; Arthropoda; Insecta; Pterygota;
 Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
 Ephydroidea; Drosophilidae; Drosophila.
 1 (bases 1 to 39839)
 ADAMS, M. and VENTER, J.C.
 Direct Submission
 Submitted (03-NOV-1999) Celera Genomics, 45 West Gude Drive,
 Rockville, MD, USA
 COMMENT This sequence was identified as CDM:10212234 by the submitter.

For further information on this sequence you may e-mail to fly@celera.com.
 * NOTE: This is a 'working draft' sequence.
 * This sequence will be replaced
 * by the finished sequence as soon as it is available and
 * the accession number will be preserved.

FEATURES
 Location/Qualifiers
 1..39839
 /organism="Drosophila melanogaster"
 /db_xref="taxon:7227"
 BASE COUNT 11426 a 9282 c 8335 g 10796 t
 ORIGIN

Query Match 17.2%; Score 981.2; DB 2; Length 39839;
 Best Local Similarity 95.9%; Pred. No. 1.7e-182;
 Matches 1053; Conservative 6; Mismatches 9; Indels 30; Gaps 4;

QY 4638 CGGTAGGCGATGAAGTGAATAATAATAAAGCGGAGCGAGTAGAGATGAAGCGAGAGAG 4697
 Db 39838 CTGTAGGCGATGAAGTGAATAATAATAAAGCGGAGCGAGTAGAGATGAAGCGAGAGAG 39779
 QY 4698 CGAGTCAGGACCTTCGACCTTTAACTGAACATAGTATATCTTGTGCACCTACTCTCCAC 4757
 Db 39778 CGAGTCAGGACCTTCGACCTTTAACTGAACATAGTATATCTTGTGCACCTACTCTCCAC 39719
 QY 4758 AACAAATATATATTTTAAATTTAGAAATTCAAAAGGCGCAACTGGAATTCGAAACCTT 4817
 Db 39718 AACAAATATATATTTTAAATTTAGAAATTCAAAAGGCGCAACTGGAATTCGAAACCTT 39659
 QY 4818 TCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4877
 Db 39658 TCTGTGCT-----CAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 39609
 QY 4878 CGAATTTACCAACCACTTCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 4937
 Db 39608 CGAATTTACCAACCACTTCACCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 39549
 QY 4938 GAACTCAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 4997
 Db 39548 GAACTCAGCAGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 39489
 QY 4998 GCGAGTGTGCGATGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCG 5057
 Db 39488 GCGAGTGTGCGATGAGAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCG 39429
 QY 5058 GTCTTATAGCCTTAATCTAAATCTAACTGGGAGACAGGACCCATGTA----- 5106
 Db 39428 GTCTTATAGCCTTAATCTAAATCTAACTGGGAGACAGGACCCATGTAATGTAATAG 39369
 QY 5107 TGTCTCTGCTATCCAAATTCGCTCTATCTATCTATCTATCTATCTATCTATCTATCTATCT 5166
 Db 39368 TGTCTCTGCTATCCAAATTCGCTCTATCTATCTATCTATCTATCTATCTATCTATCT 39309
 QY 5167 C-----CTCCCCATCCAAAAGCAAACTTAGACGTAGCCTATGTGTAAGAGCTAGCAATGT 5222
 Db 39308 CTCTCATCTCCATCCAAAAGCAAACTTAGACGTAGCCTATGTGTAAGAGCTAGCAATGT 39249
 QY 5223 TAGACCAACTTGTGAATCCAAATGAATTTGTAGCCCGCAGGAGAAAGCGCGGGA 5282
 Db 39248 TAGACCAACTTGTGAATCCAAATGAATTTGTAGCCCGCAGGAGAAAGCGCGGGA 39189
 QY 5283 AATTCAACACTTATCTCTGATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 5342
 Db 39188 AATTCAACACTTATCTCTGATAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 39134
 QY 5343 ACAGTACGAGAAATGTAAATCTTCTTAATTAATTTGTAAGAACACAGTAAATGTAA 5402
 Db 39133 ACAGTACGAGAAATGTAAATCTTCTTAATTAATTTGTAAGAACACAGTAAATGTAA 39074
 QY 5403 TCTATCTAGAGTGTGTAGCGCCCTTAAGATGTTTTTTAGTTTATAGACCGCTAACCGTA 5462
 Db 39073 TCTATCTAGAGTGTGTAGCGCCCTTAAGATGTTTTTTAGTTTATAGACCGCTAACCGTA 39014

QY 5463 ATCTAGTTTAATTCCTAACACTAAACGAGAGTACAGTACATGGTGTCTTTTCTTTTCTGCT 5522
DB 39013 ATCTAGTTTAATTCCTAACACTAAACGAGAGTACAGTACATGGTGTCTTTTCTTTTCTGCT 38954
QY 5523 AGTTCGTTGGAATAATGCTTAAACGGCAACAGATTTGTTTCTTTTCTTTTAAATAGCTTCAGT 5582
DB 38953 AGTTCGTTGGAATAATGCTTAAACGGCAACAGATTTGTTTCTTTTCTTTTAAATAGCTTCAGT 38894
QY 5583 TTGTATGTCGCTGTGTTTATTAATGACTTATATATAGTCCATCTGAATATTCGTTGATG 5642
DB 38893 TTGTATGTCGCTGTGTTTATTAATGACTTATATATAGTCCATCTGAATATTCGTTGATG 38834
QY 5643 GAGCCTATTTAAATGTCGAGTCGAGCTAAATGGAAGAAATACAAACAACTCTGTGTGC 5702
DB 38833 GAGCCTATTTAAATGTCGAGTCGAGCTAAATGGAAGAAATACAAACAACTCTGTGTGC 38774
QY 5703 CTTGGCCAAATAGTTTAC 5720
DB 38773 CTTGGCCAAATAGTTTAC 38756

RESULT 8
LOCUS AR201457 3155 bp DNA linear PAT 20-APR-2002
DEFINITION Sequence 7 from patent US 6359193.
ACCESSION AR201457
VERSION AR201457.1 GI:20252345
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 3155)
AUTHORS Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.
TITLE Nucleotide sequences of lats genes
JOURNAL Patent: US 6359193-A 7 19-MAR-2002;
FEATURES Location/Qualifiers
source 1..3155
BASE COUNT 751 a 924 c. 894 g 586 t
ORIGIN

Query Match 9.98; Score 567.8; DB 6; Length 3155;
Best Local Similarity 64.38; Pred. No.: 3.3e-101;
Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;

QY 2869 GGGCGGAATAGCTCGGGCGGAGCAACGGATCCACGGGACCAACCGCTCTCTCGTCGAC 2928
DB 1329 GGGCACTGCTAGACGGAGTGCAGAGAGCCAAAGGTCTCGAAGGAGCAAGCTGG 1388
QY 2929 CAGCTGCAAGAGATCAAGACGGCTCGCCCATCCCGAGGCGCAAGAGATCTCCAAAGGA 2988
DB 1389 CAGAGACAAAAGCAGATTTCAGACCTCCCGGTGCTGTCCGCAAGATAGC----- 1440
QY 2989 GAAGGAGGAGCGCAAGGAGTTCCGCATCAGCAGTACTCGCCGCAAGCCTTCAAGTT 3048
DB 1441 -AGAGATGAAGAGAGAGAGTCTCGATCAAGAGTTACTTCCCTTATGCTTCAAAAT 1499
QY 3049 CTTCATGAGCAGACATAGAGACGTGATCAAGTCTGATCGCCAGCGCAGCTATCGCAA 3108
DB 1500 CTTCATGAGCAACACCTGGAGATGTCTCAAAACCTACCAGCAGAGGTGACCGCGAG 1559
QY 3109 GAATCAGCTGGAGAGGATGCAAAAGTGGGACTGCCCGATCAGACCCAAATTCGAGAT 3168
DB 1560 GCTACAGCTGGAGAGGAAATGGCAAGCTCGGCTCTGTGAGGCGGAGCAGGAGCAGAT 1619
QY 3169 GAGGAAATGCTGAACCAAGAGAGCAACTACATTCGATTGAGCGCGCCAGATGGA 3228
DB 1620 GAGGAAATGCTTACCAGAGAGAGTCTAATCAACCGGCTGAAGGGCGCAAGATGGA 1679
QY 3229 CAAGAGATGTTGCTCAAACTGAAGCCCATTTGGAGTGGGTGCTTTTGGCGAGGTAAAGCT 3288
DB 1680 CAAGTCCATGTTTGTGAATCAAGACTCTAGGCATCGGTGCCCTTTGGGGAAGTGTGCT 1739

QY 3289 GGTGAGCAAAATTCGATACCTCGAACCATTTGTATGCGATGAAACCCCTGCGGAAGACGGA 3348
DB 1740 CGCTTTGTAAGCTGGA---CACTCAACGCTCTGTACGCCATGAAGACTCTCAGGAAGAAGGA 1796
QY 3349 GGTTCCTAAGCGGATCAGGTGGCACAGCTGAAGCCCGAGAGGGATATCTTCGCGGAAGC 3408
DB 1797 TGTCTTGAACCGGAATCAAGTGGCCCATGTCAAGGCTCAGAGGGACATCTCTGGCTGAAGC 1856
QY 3409 CGACAATAACTGGGTGGAAGTTGTACTACAGCTTCCAGGACAAAGGATAATCTCTACTTT 3468
DB 1857 AGACAATAGTGGGTGGAAGTTGTACTACTACTTCTTCCAGGAAGAAGACAGCTGTACTTT 1916
QY 3469 TGTGATGACTATACATACAGGTGTGTATCTGTATGCTGCTCATCAAACTGGGCATTTT 3528
DB 1917 TGTGATGACTATACATACAGGTGTGTATCTGTATGAGCTGTGATCAGGATGGAGGTCTT 1976
QY 3529 CGAGGAGAACTGGCCAGATTTCTACATCGCGAGGTACCTGCGCGGTGGACAGGTTCA 3588
DB 1977 CCCTGAGCACCTGGCCGCTTCTACATTTGCAGAGTTGACCTGGCCATTTGAAAGTGTCCA 2036
QY 3589 CAAAATGGCTTCTACACAGAGACATCAAGCCTGACAAACATCTCATCTGATAGGAGGG 3648
DB 2037 CAAGATGGCTTTTATCCACCGGACATCAAGCCTGACAAACATCTCATCTGACCTGGATGG 2096
QY 3649 ACACATAAAGCTCACCGACTTTGGCCTGTGTGACCGGATTTCCGATGGAGCGCAACTCGAA 3708
DB 2097 TCATATTAAGCTGACAGATTTTGGCCTCTGCACTGGATTAGGTGGAATCAATTCCAA 2156
QY 3709 GTACTACAGAGAAAGCGCAATCACTCGCGCCAGGACTCGATGAGGCTCTGGAGGAT- 3767
DB 2157 GTACTACC---AGAAAGGGAACCATGAGACAGGACATGAGGCGCGGTGACCTCTG 2213
QY 3768 -----ACTCCGAGAAGCGGACCAAGCCGCTGCTGGAGGCGGCACGGAT 3813
DB 2214 GGACGATGTTTCCAACTGTCTGCTGTGGAGACAGGTAAAGACCTTGGAGCAGAGGGCGCA 2273
QY 3814 GCGGATCACAAAGAGTCTTGGCCCACTCGCTGTGGGACACCCCGAACTACATAGTCC 3873
DB 2274 GAAGCAGCACCAGAGGTGCTTGGCAGCATTTCTTGTGGGACACCAAAATTCATCGCTCC 2333
QY 3874 CGAGGTCTGGAGAGAGTGGGTACACGACGCTGTGCGACTCTGGAGCGTGGGCGTCTAT 3933
DB 2334 GGAGGTCTTCTCGCAAAAGGTTACACGACGCTGTGACTGTGGTGGAGCGTCTGAT 2393
QY 3934 CCTTATGAGATGTGTGGTTCAGCGCCCTTCTTGGCCAAACAGTCCGCTGGAAGACGA 3993
DB 2394 TCTCTTTGAGATGTGTGGTGGGACGCGCTTCTTGGCCCCACCCCGCAGAGACGCA 2453
QY 3994 ACAAAAGTTCATCAACTGGGAGAAACCTGTCATATTTCCGCGCGAGCGCGAGTTATCCG 4053
DB 2454 GCTGAAGGTGATCAACTGGGAGAGCAGCTGCATATCCCTACGAGGTGAGGCTCAGCGC 2513
QY 4054 CGAGGCTACGACTTGTATAGGAGGCTCTGTGCTCGCTGCAAGGCTGCAAGGCGGTGGCA ---A 4110
DB 2514 TGAGGCGCGAGACCTCATCAGAAAGCTGTGCTGCGCGGCTGACTGCGCGCTGGCGAGGA 2573
QY 4111 GAGCGCTGAGCAGGTCAAGAGCCACGACTTCTTCAAGGGCATCGACTTTGC---GGACAT 4167
DB 2574 TGGGCGAGATGACCTCAAGGCACACCCGTTCTTCAACACCATCGACTTTTCCCGTGCAT 2633
QY 4168 GCGGAAGCAGAAAGCGCCCTTACATACCGGAAATCAAGCAACCCACCGACACATCCAACTT 4227
DB 2634 CCGAAAGCAGGCTGCACCCCTACGTCGCCACCATCAGCCACCCCATGACACCTCCAATTT 2693
QY 4228 TGATCCCGTGGATCCGAGAGGCTGCGCTCGATGACTTCCACATGAGCAGCGCGCATGA 4287
DB 2694 TGACCCCGTGGATGAAGAAAGCCCTTGGCACGAGGCGCGGAGAGAGAGCCCAAGGCGCTG 2753
QY 4288 TGTGACCAAGATGACCGCACATTTCTTCAAGGGCATCGACTTTGC---CAGCGCTTTTTCGAATTTACCTT 4335
DB 2754 GGACAGCTGGCTCCCGCAGCAGCAGATCCAGACGACGCTTCTATGNGTTCACCTT 2813
QY 4336 CCGTGGCTTCTTCGACGACAA 4356


```

Db 2814 CCGCAGGTTCTTCGATGACAA 2834
|||||
AB0233958
Mus musculus MmLATS2 mRNA for warts/lats- like kinase, complete cds.
DEFINITION
ACCESSION AB0233958
VERSION 1.0
KEYWORDS MmLATS2; warts/lats-like kinase.
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (sites)
AUTHORS Yabuta,N., Fujii,T., Copeland,N.G., Gilbert,D.J., Jenkins,N.A.,
Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
Nishiguchi,H., Endo,Y., Toji,S., Tanaka,H., Nishimune,Y. and
Nojima,H.
TITLE Structure, expression, and chromosome mapping of LATS2, a mammalian
homologue of the Drosophila tumor suppressor gene lats/warts
JOURNAL Genomics 63 (2), 263-270 (2000)
MEDLINE 20139436
REFERENCE 2 (bases 1 to 3460)
AUTHORS Nojima,H. and Fujii,T.
TITLE Direct Submission
JOURNAL Submitted (22-FEB-1999) Hiroshi Nojima, Research Institute for
Microbial Diseases, Osaka University, Department of Molecular
Genetics; Yamadaoka 3-1, Suita, Osaka 565-0871, Japan
(E-mail:hnojima@biken.osaka-u.ac.jp, Tel:81-6-6875-3980,
Fax:81-6-6875-5192)
COMMENT Sequence updated (06-Jan-2000).
FEATURES
Location/Qualifiers
1..3460
/organism="Mus musculus"
/db_xref="taxon:10090"
1..3460
/gene="MmLATS2"
134..3262
/gene="MmLATS2"
/feature="WALR"
/codon_start=1
/product="warts/lats-like kinase"
/protein_id="BAA92380.1"
/db_xref="GI:7212786"
/translation="MRPKTEPATYSGNSRQRLQETREGLKQPSKASTOGLLVGPNSD
TSLDAKVLGSDASRQQRATPKQYKALRETRYSLPPANESGTSAAEYVNRQM
LOELVAGCDQMGARLQKQSRSEAEALYISKMGYLDPRNEQIVRVIKQTSPEGK
LAPVTRRPSFEGTGEALPSYHRLGGNAYEGPALEMPROYLDLFPAGAGTHGA
QAQHPKPGYSTAVPSAHFPGTHYGRHLLSEQPGYGVQRSSFPQNTPPDAYSSMA
KAQGGPASPFPAGHLYTASHHKPAATPPGAPHLVLTGRPTFTGESSAOAVLAP
SRNSLNADLYELGTVPWSAAPLABRDSLOKQGLEASRPHVAFRAGPSRTSNFNPQ
EESGAPHLVDVYGGSSRRPPPKHLLLPSSKSEQYSDVLSICTVQOSLRGTE
QDRSKHGAKQKAGRDKKQIQTSPPVNRNDRDEKRESIKSYSPYAFKPFMEQ
HVENYIKYQKSRRLQEQEVAKAGLCEAQEQMRKILYOKESYNRKLRKAKMDKS
MFVKTKTGAGFGVCLACKLDTALYAMKTLRKQVLRNQNVAHVAERDILAEAD
NEWVYLYSKQSDSLYEVADYIPGQDMSSLLIRNEVEPEHLAFYIAELTLAESV
HKMGFIHDKIPDNLIDLGHKILDTGLCTGFRWTHNSKTYQGNHNRQSMPEPG
LWDDVNSCRGDRLTLEORAKQHORCLASLSVTPNYIAPEVLLRKYQLCLDWNS
VGVLIFMLVGQPPPLPTPTQKLVNWSLETHTPTQVRLSABARDLITKLCQAD
CRLRGDADMDLKAHPFNTIDFSRDIRKOPAPVPTISHPMDTSNFDVDESPWHEA
SGESAKMDTLASPSKHPHAHYEFTFRFRDDNGYFPCPKPEPAESADPGDADL
EGAEAGCQPVVY"
3460
polyA_site
BASE COUNT 861 a 992 c 963 g 644 t
ORIGIN
Query Match 9.9%; Score 565.4; DB 10; Length 3460;
Best Local Similarity 65.3%; Pred. No. 9.8e-101;
Matches 954; Conservative 4; Mismatches 455; Indels 48; Gaps 7;

```


QY 4012 GGAGAAACGCTCATATTCCGCGCAGGCGGAGTTATCCGCGAGGCTACGGACTGTAT 4071
 Db 3170 GGAGAACACGCTCCACATCCAGCCAGGTGAAGCTGAGCCCTGAGGCCAGGACCTCAT 3229
 QY 4072 AAGGAGGCTCTGTGCGCTGACAAAGCGGCTGGGCAAGAGCGTGGACAGG---TCAA 4128
 Db 3230 CACCAAGCTGTGTGCTCCGACAGCACCGCCCTGGGCGGAATGGGCGCGATGACCTGAA 3289
 QY 4129 GAGCCACGACTTCTTCAAGGGCATCGACTTTGC---GGACATCGGGAAGCAGAAAGCGCC 4185
 Db 3290 GGCCACCCCTCTTTCAGCGCAATGACTTCTCAGTGACATCCGGAAGCAGCAGCCGCC 3349
 QY 4186 CTATACACCGAAATCAAGACACCCACGACATCAACTTTGATCCGCTGGATCCGGA 4245
 Db 3350 CTACGTTCCCAACCATCAGCCACCCATGGACACCTCGAATTTCCGACCCCGTAGATGA 3409
 QY 4245 GAGCTGCGCTCGAAGTCACTCCACCATGACGACGCGGATGATGTC-----GA 4293
 Db 3410 AGCCCTTGAAGCATGCCAGCGAAGGTAGCACCAAGGCTGGGACACACTCACCTCGCC 3469
 QY 4294 CAGAATGACCGCAGCTTCCACCGCTTTTTCGAATTTACCTTCGCTCGCTTCTTCGACGA 4353
 Db 3470 CATACACAGCATCTCTGACGACGCAATTTACGAATTCACCTTCCGAAGGTTCTTTGATGA 3529
 QY 4354 CAA 4356
 Db 3530 CAA 3532

RESULT 11

E38227 Human tumor regulatory gene. 5486 bp DNA linear PAT 31-JAN-2002
 DEFINITION Human tumor regulatory gene.
 ACCESSION E38227
 VERSION E38227.1 GI:18626936
 KEYWORDS JP 2000210086-A/2.
 SOURCE Homo sapiens.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 1 (bases 1 to 5486)
 Koga, J., Kono, K. and N, Z.F.
 Human tumor regulatory gene
 Patent: JP 2000210086-A 2 02-AUG-2000;
 NIHON CHEMICAL RESEARCH K K
 OS Homo sapiens (human)
 PN JP 2000210086-A/2
 PD 02-AUG-2000
 PF 25-JAN-1999 JP 1999016223
 PR JUNICHI KOGA, KEIKO KONO, ZOROTARYO FUJIDORI N PC
 C12N15/09, A61K31/00, A61K38/27, C07K14/47, C07K14/525, C12Q1/68, PC
 G01N33/50,
 PC G01N33/50, G01N33/574, G01N33/577//C12P21/08, (C12N15/09, PC
 C12R1:91), C12N15/00,
 PC A61K37/36, (C12N15/00, C12R1:91)
 CC CC
 FH Key Location/Qualifiers
 FT Source 1..5486
 FT Location/Qualifiers
 /organism='Homo sapiens'
 /db_xref='taxon:9606'
 1..5486

FEATURES

BASE-COUNT 1380 a 1451 c 1388 g 1267 t
 ORIGIN
 Query Match 9.8%; Score 558.6; DB 6; Length 5486;
 Best Local Similarity 64.1%; Pred. No. 2.3e-99;
 Matches 964; Conservative 4; Mismatches 493; Indels 42; Gaps 7;
 QY 2884 GGGCGCAGCAAGGATCCACCGGACCAACCGCCTCTCTGTCAGCAGCTGCAAGAAGAT 2943
 Db 2884 GGGCGCAGCGCCCTTTTGGCACCTACTCCACAGAAACCCAGCTGAAGGTGATCAACTG 3169

Db 2042 GGCGCGCAGCAGAGCGCGCAAAAGCGCCAAAGGGGAGCAAAAGGATAAAAAAGCA 2101
 QY 2944 CAGCAGCGCTCCCGCATCCCGAGCGCAGAGAAGATCTCCAAGGAAGAGGAGGAGCG 3003
 Db 2102 GATTCAGACCTCTCCCGTTCGCGTCCGCAAAACAGC-----AGAGACGAAGAGAA 2152
 QY 3004 CAAGGAGTTCCCGCATCAGGCGAGTACTCGCGCAAGCCTTCAAGTTCTTCTTCAAGTCAAGCAGCA 3063
 Db 2153 GAGAGAGTCACGATCAAGAGCTACTCGCCATACGCTTTAAGTTCTTCTTCAAGGAGCA 2212
 QY 3064 CATAGAGACGTGATCAAGTCGTATCGCAGCGCAGCTATCGCAAGAAATCAGCTGGAGAA 3123
 Db 2213 CTTGGAGATGTCTAAAACCTTACCAGCAGAGGTTAAACCGGAGCTGCAGCTGGAGCA 2272
 QY 3124 GGAGATGCACAAAGTGGGACTGCCCATCAGACCCAAATCGAGATCAGGAAATGCTGAA 3183
 Db 2273 AGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATCGGGAAGATCTCTTA 2332
 QY 3184 CCAAAAGGAGAGCAACTACATTCGATTGAAGCGCGCAAGATGGACAAGAGCATGTTCTGT 3243
 Db 2333 CCAGAAAGAGTCTAATTACAACAGGTTAAGAGGGCCAGATGGACAAGTCTATGTTGT 2392
 QY 3244 CAAACTGAAGCCCATTTGGAGTGGTGCATTTGGCGAGGTAACGCTGGTGGAGCAAAATCGA 3303
 Db 2393 CAAAGTCAAAACCCCTGGGATCGCTTGGAGAAAGTGTGCTTGTGTAAGGTGA 2452
 QY 3304 TACCTCGAACCATTTGATCGGATGAAACCCCTCGGAAAGCGGAGCTTCTCAAGCGGAA 3363
 Db 2453 ---CACTCACGCGCTGTACGCCATGAAGACCCCTAAGGAAAAAGGATGTCTGAAACCGGAA 2509
 QY 3364 TCAGTGGGCACACGTCAAGCGCGCAGAGGATATCTTCGCGGAAGCGGCAAAATCAACTGGGT 3423
 Db 2510 TCAGTGGCGCCAGTCAAGCGCGCAGAGGACATCTCTGGCGGAGCAGCAATCAGTGGGT 2569
 QY 3424 GTGAAGTTGTACTACAGCTTCAGGACAAGGATAATCTGTACTTTGTGATGGACTACAT 3483
 Db 2570 GGTCAAACTCTACTACTCTCTTCCAGACAAAGACAGCTGTACTTTGTGATGGACTACAT 2629
 QY 3484 ACCAGTGTGATCTGATCTGCTGCTCATCAAACTGGGCATTTTCGAGAGGAACTGGC 3543
 Db 2630 CCTGTGGGGACATGATGAGCTGCTGATCGGATGGAGGTCTTCCCTGAGCACCTGGC 2689
 QY 3544 CAGATTCTACATCGCGCAGGTCACTTCGCGCGTGGAGCGTTTCACAAAATGGGCTTCAT 3603
 Db 2690 CCGTTCTACATCGCAGAGTGTGCTTGGCCATTGAGAGTGTCCACAAAGTGGGCTTCAT 2749
 QY 3604 TCACAGACATCAAGCCTGACAACTACTCATCATAGTAGGAGCGGAGCAGACATAAGCTCAC 3663
 Db 2750 CCACCGAGACATCAAGCCTGATAAATTTTGTAGATCTGGATGGTTCACATTAACCTCAC 2809
 QY 3664 CGACTTTGGCCTGTGCACGGGATTCGATGGAGCGACACACTCAAGTACTACAGGAGA- 3722
 Db 2810 AGATTTCGGCCCTCTGCACCTGGGTTTCAGTGGGACTCAAAATCCAAATATATACCAGAAAG 2869
 QY 3723 -----ACGGCAATCACTCGCGCAGGACTCGATGGAGCCCTGGGAGGAAATCTCCGA--- 3774
 Db 2870 GAGCCATGTACAGCAGGACAGCATGAGCCAGCAGCTCTGGATGATGTGTCTCACTG 2929
 QY 3775 ---GAACGGACCGAAGCCCGTGTGGAGGCGGAGCGGATGCGCGATCACCAGAGAT 3831
 Db 2930 TCGGTGTGGGACAGGCTGAAGACCTTAGAGCAGAGGGCGGGAAGCAGCAGAGGTG 2989
 QY 3832 CTGGCCCACTCGCTGGTGGGACCCCGCACTACATAGTCTCCGAGGTCTCTGGAGAGAG 3891
 Db 2990 CCTGGCATTCTCTGGTGGGACTCCAACATACATCGCACCCGAGGTCTCTCTCGGAA 3049
 QY 3892 TGGGTACGCGAGCTCTGCGACTTACTGGAGCGTGGCGCTCATCTTATYAGATGCTGTG 3951
 Db 3050 AGGTACACTCACTGTGACTGTGGAGTGTGGAGTGTCTCTTCGAGATGCTGTG 3109
 QY 3952 GGGTCAGCGCCCTTTCTGGCCCAAGTCCCGTGGAAACCGCAACAAAGGTTCATCAACTG 4011
 Db 3110 GGGGAGCGCGCCCTTTTGGCACCTACTCCACAGAAACCCAGCTGAAGGTGATCAACTG 3169

[illegible]

QY	3304	TAC	TCGAA	CAATTG	TATCG	GATGAA	AACCC	TGCG	GAAGCG	AGCTT	CTCA	AGCG	GAA	3363
DB	2441	---	CAC	TGCG	CCCTG	TAC	CCCAT	GA	AGAC	CCCT	AGG	AAAA	AGGATG	3423
QY	3364	TC	AGGT	GGC	AC	AGT	GA	AGC	CG	GAG	AGG	GATAT	CTC	3423
DB	2498	TC	AGGT	GGT	CC	AGC	TC	CA	AGC	CG	GAG	AGG	GACAT	3423
QY	3424	GGT	GAA	CTT	G	TACT	AC	AGC	TT	CC	AG	GAC	AA	2557
DB	2558	GGT	CA	AACT	CT	ACT	TC	CTT	CC	AG	CA	AA	AG	3483
QY	3484	AC	AGT	GTG	T	G	T	G	T	G	T	G	T	3483
DB	2618	CC	T	GTG	GG	G	AG	CA	T	G	A	T	G	3483
QY	3544	C	A	G	A	T	T	C	C	G	G	A	T	3603
DB	2678	CC	GT	T	C	A	T	C	A	T	C	A	T	3603
QY	3604	T	C	A	G	A	G	A	C	T	A	G	C	3663
DB	2738	CC	A	C	G	A	G	A	C	T	A	G	C	3663
QY	3664	C	G	A	T	T	C	C	G	G	A	T	T	3722
DB	2798	A	G	A	T	T	C	C	G	G	A	T	T	3722
QY	3723	---	AC	G	CA	T	A	C	T	C	G	C	A	3774
DB	2858	G	A	G	C	A	T	G	C	A	G	A	G	3774
QY	3775	---	G	A	C	G	C	A	G	C	A	G	C	3831
DB	2918	T	C	G	G	T	G	G	G	A	C	G	G	3831
QY	3832	C	T	G	C	C	A	T	C	T	G	G	C	3891
DB	2978	C	T	G	C	C	A	T	C	T	G	G	C	3891
QY	3892	T	G	G	T	A	C	G	A	G	T	G	C	3951
DB	3038	A	G	G	T	A	C	A	T	C	A	T	C	3951
QY	3952	G	G	G	T	C	A	A	C	A	G	C	T	4011
DB	3098	G	G	G	T	C	A	A	C	A	G	C	T	4011
QY	4012	G	G	A	A	A	C	T	G	A	T	T	C	4071
DB	3158	G	G	A	A	A	C	T	G	A	T	T	C	4071
QY	4072	A	A	G	A	G	T	G	T	G	C	G	T	4185
DB	3218	C	A	C	A	A	G	T	G	T	G	C	G	4185
QY	4129	G	A	G	C	A	G	T	T	C	A	A	G	4245
DB	3278	G	G	C	C	A	C	C	T	T	C	A	A	4245
QY	4186	C	T	A	C	A	T	A	C	C	C	A	T	4393
DB	3338	C	T	A	C	T	T	C	A	A	G	T	G	4393
QY	4246	G	A	G	C	T	G	A	T	G	A	T	G	4423
DB	3398	A	G	C	T	G	A	T	G	A	T	G	A	4423
QY	4294	C	G	A	A	T	G	C	C	G	C	T	T	4533
DB	3458	C	A	T	A	C	A	G	A	T	T	A	G	4533

QY	4354	CAA	4356
Db	3518	CAA	3520
RESULT 14			
AR201456			
LOCUS	AR201456		
DEFINITION	Sequence 5 from patent US 6359193.	DNA	linear PAT 20-APR-2002
ACCESSION	AR201456		
VERSION	AR201456.1	GI:20252344	
KEYWORDS			
SOURCE	Unknown.		
ORGANISM	Unknown.		
REFERENCE	Unclassified.		
AUTHORS	Xu, T., Tao, W., Wang, W., Zhang, S. and Yu, W.		
TITLE	Nucleotide sequences of lats genes		
JOURNAL	Patent: US 6359193-A 5 19-MAR-2002;		
FEATURES	Location/Qualifiers		
source	1..3213		
BASE COUNT	946 a 764 c 696 g 807 t		
ORIGIN			
Query Match 9.1%; Score 522.8; DB 6; Length 3213;			
Best Local Similarity 65.3%; Pred. No. 2.4e-92;			
Matches 850; Conservative 4; Mismatches 421; Indels 27; Gaps 5;			
QY	2984	AAGGAGAAGGAGGAGGCCGAAGGAGTTCGGCATCAGGCAGTAGCTCGCCGCAAGCGCTTC	3043
Db	1354	AACAAGAAAGATGAAGAACGAAGAGACTCTCGGATTTCAGAGTTTACTCCCACAGGCGCTTT	1413
QY	3044	AAGTCTCTTCATGGAGCAGCACATAGAGAAGCTGATCAAGTCGTATCGCCAGCGCAGCTAT	3103
Db	1414	AAGTCTCTTCATGGAGCAGCAGTAGAAGCGTCTGAGTCTCATCAGCAGCGCTCTGCAT	1473
QY	3104	CGCAAGATCAGCTGGAGAGAGGAGATGCCAAGAATGGGACTGCCGATCATAGACCCCAATC	3163
Db	1474	CGGAAGAAGCAGCTAGAAAAATGAAATGATCGGGTTGGATTATCTCAAGATGCCAGGAT	1533
QY	3164	GAGATGAGGAAAATGCTTGAAACC AAAAGAGAGACAATCATATTCGATTGAGCCGCGCAAG	3223
Db	1534	CAATGAGAAGATGCTTTC CCGAGAAAGACTCTAATCTATTCGTCTTAAAGGGCTAAA	1593
QY	3224	ATGGACAAGACATGTTTGGTC AAATGAAGCCCATTCGAGTGGGTGCATTTGGCGAGGTA	3283
Db	1594	ATGGACAAGTCTATGTTTGTGTAAGATAAAGACATTAGGAATAGGACGCTTGGTGAAGTC	1653
QY	3284	ACGCTGCTGAGCAAAATCGATACCTCGAACCATTTGTATGCCATGAAACCCCTGGGAAA	3343
Db	1654	TGCTAGCAGAAAGTCGATA ---CTAAGCTTTGTATGCAACAAAGACTCTTCGAAG	1710
QY	3344	CGCGAGCTTCTCAAGCGGAATCAGGTGGGCACACGTTGAAGCCGCGAGAGGATATCTCCG	3403
Db	1711	AAAGACGTTCTGCTCCGAAATCAGGTGGCTCATGTCAAAGCGGAGGATATCTACGA	1770
QY	3404	GAGCCGCAAAATAACTGGTGGTGTGAAGTTGTACTACAGCTTCCAGGACAAGGATATCTG	3463
Db	1771	GAGCCGCAAAATAAGTGGGTGGTTCGCGCTGTACTACTCTTCTCCAGGACAAGACAATTC	1830
QY	3464	TACTTTGTATGGACTACATACAGGTGGTGTATCTGATCTCCCTGCTCATCAAACTGGGC	3523
Db	1831	TACTTTGTATGGACTACATCTCTGGGGGGATATGATGAGCCTATTAATTAGAAATGGGC	1890
QY	3524	ATTTTCGAGGAGGAATCGGCCAATTTCTACATCGCGGAGGTCACTGCGCGCTGGACAGC	3583
Db	1891	ATCTTTCTTGAAAAATCTGGCAGGATTTACATAGCAGACTTACCTGTGCGAGTTGAAGT	1950
QY	3584	GTTCAAAATATGGGCTTCAATTCACAGAGACATCAAGCCTGACAACTACTCATCATAGG	3643
Db	1951	GTTCAAAATATGGGTTTTATTTCATAGAGATATTAACCTGATCAATTTTCATTGACCGT	2010

[illegible]

QY 3344 GCGGAGCTTCTCAAGCGGAATCAGGTGGCACACGTGAAGCGGAGGAGGATATCTCGCG 3403
 Db 1711 AAAGAGCTTCTGCTCGGAATCAGGTGGCTCATGTGAAGCGGAGGAGGATATCTAGCA 1770
 QY 3404 GAAGCGGACAAATCACTGGTGAAGTTGTACTACACCTTCCAGGACAAGGATAATCTG 3463
 Db 1771 GAAGCGGACAAATCACTGGTGGCTCGCTACTACTCTTCCAGGACAAGGACAACCTG 1830
 QY 3464 TACTTTGTGATGACTACATACACGCTGGTGTGATCTGTGATCTGCTCATCAAACTGGCG 3523
 Db 1831 TACTTTGTGATGACTACATCTTGGGGGATATGATGACCTTAAATAGAACTGGC 1890
 QY 3524 ATTTTCGAGGAGGAGTGGCGAGATTTACATCGCCGAGGTCACTCGCGCTGGACAGC 3583
 Db 1891 ATCTTCTGAAATCTGGCAGGATTTACATAGAGATTAACCTGATACATTTTGAATGACCGT 1950
 QY 3584 GTTCACAAATGGGCTTCATTACAGAGACATCAAGCTTCCGATGGACGACACAC 3703
 Db 1951 GTTCATAAATGGGTTTATTCATAGAGATTAACCTGATACATTTTGAATGACCGT 2010
 QY 3644 GACGACACATTAAGCTACCGACTTTGGCTGTGACGGGATTCGATGGACGACACAC 3703
 Db 2011 GATGCCATATTAATTAAGTACTGACTTTGGCTTGTGACCTTGTGATGGACACATGAC 2070
 QY 3704 TCGAAGTACTACAGGAGAGAGGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAG 3763
 Db 2071 TCGAAGTACTACCGGAGGATTCACCCAGGAGATAGCATGGATTTTCACTAGTAC 2127
 QY 3764 GAATACTCCGA-----GAACGGACCGAAGCCACCGCTGTGGAGAGGCGA 3808
 Db 2128 GAATGGGAGACATCTTCCAAATTCGCTGTGGGACAGACTGAAGCCACTGGAGCGGAG 2187
 QY 3809 CGGATCGCGATCACCAAGAGTCTGGCCACTCGCTGTGGGACCCCGAAGTACATA 3868
 Db 2188 GCTGCTCGCGAGCAGGAGTCTAGCCATCTCTGTTGGGACTCCCAATTAAT 2247
 QY 3869 GCTCCGAGGTCTGAGAGAGTGGGTACAGCAGCTGTGCGACTACTGGAGCGTGGC 3928
 Db 2248 GCACCTGAAGTCTACTGCGACAGATATACACAGCTGTGACTGTGGAGTGTGGT 2307
 QY 3929 GTCTCTCTTATGAGAGTCTGGGTACGCGCCCTTTCTGGCCAAACAGTCCGCTGGAA 3988
 Db 2308 GTATTTCTTTGTGAATGTGGTGGGACAACTCTCTTGGCACAAACCCATTAGAA 2367
 QY 3989 AGCAACAAAGTCACTCACTGGGAGAAACSTGATATTTCCGCGCAGGCGGAGTTA 4048
 Db 2368 ACACAAATCAAGTTATCATCTGGCAAACTTCTTACACATCCCTCTCAAGCTAAGCTG 2427
 QY 4049 TCCCGGAGGTACGACTTGTATAGGAGGCTCTGTGCGCTGACAGCGGCTGGC 4108
 Db 2428 AGTCTGAAGCTCTGACCTATTATCAAACTGTGTCAGGACAGAGACCGCTCGGC 2487
 QY 4109 AAGA---CGGTGGAGGCTCAAGGCCACGACTTCTTCAAGGCGATCGA---CTTTGGC 4162
 Db 2488 AAGAAGGTGTGATGAGATAAGGCTCATCCATTTTAAAGCCATCGATTTCTCTAGT 2547
 QY 4163 GACATGCGGAAGAGAAAGCCCTACATACCGGAATCAAGACCCACGAGGACATCC 4222
 Db 2548 GATCTGAGACAGCAGCTGTCTTCAATCCCTTAAATCAGGATCCACAGATACATCC 2607
 QY 4223 AACTTTGATCCCGTGGATCCGAGAACTGGCTCGAATGAC 4264
 Db 2608 AATTGACCCCTGTGATCTGATAAATTTGGAGCGATGGC 2649

GenCore version 5.1.3
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:50:07 ; Search time 729.592 Seconds
(without alignments)
17655.667 Million cell updates/sec.

Title: US-09-763-334-7

Perfect score: 5720

Sequence: 1 attagcagcagcgagcaaa.....gccttgccaattagttac 5720

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 112599159 residues

Total number of hits satisfying chosen parameters: 4370478

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

N_Geneseq_101002.*

1: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1980.DAT.*
2: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1981.DAT.*
3: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1982.DAT.*
4: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1983.DAT.*
5: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1984.DAT.*
6: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1985.DAT.*
7: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1986.DAT.*
8: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1987.DAT.*
9: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1988.DAT.*
10: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1989.DAT.*
11: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1990.DAT.*
12: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1991.DAT.*
13: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1992.DAT.*
14: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1993.DAT.*
15: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1994.DAT.*
16: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1995.DAT.*
17: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1996.DAT.*
18: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1997.DAT.*
19: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1998.DAT.*
20: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA1999.DAT.*
21: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2000.DAT.*
22: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001A.DAT.*
23: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2001B.DAT.*
24: /SIDS2/gcgdata/geneseq/geneseq-emb1/NA2002.DAT.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	ID	Description
1	5708	99.8	5720	21	AAZ51508 Drosophila melanog
2	5693.6	99.5	5720	17	AAZ42117 Drosophila melanog
3	3628	63.4	11187	23	ABL03168 Drosophila melanog
4	3242.2	56.7	3319	23	ABL03169 Drosophila melanog
5	567.8	9.9	3155	17	AAZ42120 M-lats2 gene encod
6	567.8	9.9	3155	21	AAZ51507 Mouse Lats2 (large
7	558.6	9.8	5276	20	AAZ87397 Human WART1 cDNA
8	558.6	9.8	5486	21	AAA59129 DNA encoding a tum
9	558.6	9.8	5486	21	AAA59130 DNA encoding a tum

10	557	9.7	1961	21	AAZ61160	CDNA SSI1771A encod
11	557	9.7	3533	22	AAZ17165	Human cancer-inhib
12	536	9.4	1498	21	AAZ61158	CDNA SSI1771 encodi
13	535.2	9.4	1912	23	ABK43457	DNA encoding novel
14	522.8	9.1	3213	17	AAZ42119	M-lats gene encodi
15	522.8	9.1	3213	21	AAZ51506	Mouse Lats (large
16	487.8	8.5	2043	22	AAZ03989	Human protein tyro
17	470.6	8.2	3984	17	AAZ42118	H-lats gene encodi
18	470.6	8.2	3984	21	AAZ51505	Human Lats (large
19	470.6	8.2	7382	20	AAZ87396	Human WART1 cDNA
20	469	8.2	2442	20	AAZ32981	Human Warts protei
21	455.8	8.0	1374	20	AAZ32982	Human Warts gene f
22	358.6	6.3	1357	22	ABA08740	Human large tumour
23	340.4	6.0	1501	22	AAZ75341	Human TGF-beta rec
24	321.4	5.6	676	22	AAZ27179	CDNA encoding nove
25	321.4	5.6	676	23	ABK43776	DNA encoding novel
26	290.8	5.1	2255	23	ABL29695	Drosophila melanog
27	286.8	5.0	2765	23	ABL02442	Drosophila melanog
28	273.6	4.8	638	21	AAZ79942	Human colon cancer
29	261.2	4.6	2101	17	AAZ31451	Drosophila Ndr ser
30	254.8	4.5	2160	22	AAZ66703	A. gossypii AG007
31	232.6	4.1	3018	17	AAZ31452	Human Ndr serine/t
32	232.6	4.1	3018	24	ABK84773	Human NDR differe
33	232.6	4.1	3800	22	ABA08668	Human NDR homology
34	216.6	3.8	1935	20	AAZ06834	Disease associated
35	215.4	3.8	1710	21	AAZ43403	Arabidopsis thalia
36	215.4	3.8	2003	21	AAZ49841	Arabidopsis thalia
37	215	3.8	4983	22	AAZ44629	Novel protein kina
38	204.2	3.6	1894	21	AAZ39567	Arabidopsis thalia
39	204.2	3.6	2001	21	AAZ36174	Arabidopsis thalia
40	162.6	3.5	1922	21	AAZ49426	Arabidopsis thalia
41	164.6	2.9	678	21	AAZ79982	Human colon cancer
42	161.6	2.8	734	23	ABK43709	DNA encoding novel
43	161.2	2.8	734	22	AAZ27161	CDNA encoding nove
44	161.2	2.8	734	22	AAZ56728	Human immune/haema
45	161.2	2.8	734	23	ABK43985	DNA encoding novel

ALIGNMENTS

RESULT 1

AAZ51508

ID AAZ51508 standard; DNA; 5720 BP.

XX AAZ51508;

AC AAZ51508;

XX 21-JUN-2000 (first entry)

DT Drosophila melanogaster Lats (large tumour suppressor) DNA.

DE Fruit fly; Lats: large tumour suppressor; cytostatic; vulnary;
KW cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
KW treatment; prevention; screening; cancer; skin; ovarian tumour;
KW soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
KW LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
KW dysplasia; degenerative disorder; growth deficiency; physical trauma;
KW hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

XX Drosophila melanogaster.

OS Drosophila melanogaster.

XX Key Location/Qualifiers

FT CDS 1103..4402

FT /*tag= a

FT /product= "Lats protein"

FT polyA_signal 4655..4660

FT /*tag= b

FT misc_feature 5013..5142

FT /*tag= c

FT /note= "this region is identical to the 1-141

FT nucleotides of Drosophila plc-21 transcript"

XX WO200010602-A1.

PN

XX PD 02-MAR-2000.
 XX PF 18-AUG-1999; 99WO-US19068.
 XX PR 18-AUG-1998; 98US-0096996.
 XX PR 18-AUG-1998; 98US-0096997.
 XX PA (UYUA) UNIV YALE.
 XX PI Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;
 PI Turenchalk GS, Stewart RA;
 XX DR WPI: 2000-246496/21.
 XX DR P-PSDB; AAY70393.
 XX PT Use of lats proteins, complexes of lats and cdc2 for treating cancer
 PT that is refractory to treatment by standard chemotherapy and radiation
 PT therapy, and disorders associated with aberrant levels of cdc2 activity
 XX
 XX Claim 44; Fig 15; 134pp; English.
 CC The present sequence is a DNA encoding Drosophila lats (large tumour
 CC suppressor) protein which is a cell overproliferation inhibitor and a
 CC negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
 CC The present sequence is useful for treating cancer that is refractory
 CC to standard chemotherapy or radiation therapy such as hyperplasia,
 CC metaplasia, or dysplasia, and disorders associated with aberrant
 CC levels of cdc2 activity. Conditions treated by promoting cdc2 function
 CC include degenerative disorders, growth deficiencies, hypoproliferative
 CC disorders, physical trauma, lesions, and wounds. An animal model
 CC recombination, e.g. a lats knock-out mouse, is used for screening
 CC compounds that can be used to treat or prevent cancer, particularly
 CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
 XX
 XX Sequence 5720 BP; 1684 A; 1491 C; 1457 G; 1088 T; 0 other;

Query Match 99.8%; Score 5708; DB 21; Length 5720;
 Best Local Similarity 99.5%; Pred. No.0;
 Matches 5693; Conservative 26; Mismatches 1; Indels 0; Gaps 0;

QY 1 ATCTAGCAGCGGCGGAGCAACAAACACGAAATTAATTTTACTAAATTAAGCCAAAGCG 60
 DB 1 ATCTAGCAGCGGCGGAGCAACAAACACGAAATTAATTTTACTAAATTAAGCCAAAGCG 60
 QY 61 GCATCGGAATGCCCTGAAATCGGATTCGATTCAGCGGAAAGTGGTGGCAAGCG 120
 DB 61 GCATCGGAATGCCCTGAAATCGGATTCGATTCAGCGGAAAGTGGTGGCAAGCG 120
 QY 121 GAGTGAATCAAGTGAATACGTCGGCAAAATATCAGCGAAATTCGTCGAAAGCAAGGA 180
 DB 121 GAGTGAATCAAGTGAATACGTCGGCAAAATATCAGCGAAATTCGTCGAAAGCAAGGA 180
 QY 181 AAACGGAG 240
 DB 181 AAACGGAG 240
 QY 241 GAACAAAG 300
 DB 241 GAACAAAG 300
 QY 301 AGCTCCAAAGTGGCGGCGCATCGATTCGTCGTAGGATCAATTAAGATTCGAGTGGTCGA 360
 DB 301 AGCTCCAAAGTGGCGGCGCATCGATTCGTCGTAGGATCAATTAAGATTCGAGTGGTCGA 360
 QY 361 GAATCGGCTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420
 DB 361 GAATCGGCTCAATCAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAATTAAT 420

QY 421 ATTCATTATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 DB 421 ATTCATTATCGCTCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 480
 QY 481 GTCTCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 540
 DB 481 GTCTCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCTCGGCT 540
 QY 541 AATACAAATACAAATACAAATACAAATACAAATACAAATACAAATACAAATACAAATACAAAT 600
 DB 541 AATACAAATACAAATACAAATACAAATACAAATACAAATACAAATACAAATACAAATACAAAT 600
 QY 601 GCGAAAGGGGGGTTTCTTCTTATAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 660
 DB 601 GCGAAAGGGGGGTTTCTTCTTATAATGCAATGCAATGCAATGCAATGCAATGCAATGCAATGCA 660
 QY 661 GTGGAAGTATTCGCGCGGCTGATATAATATATATATATATATATATATATATATATATATAT 720
 DB 661 GTGGAAGTATTCGCGCGGCTGATATAATATATATATATATATATATATATATATATATATAT 720
 QY 721 ACCGCGAGAAAG 780
 DB 721 ACCGCGAGAAAG 780
 QY 781 GAGATAATAAGCATCTTAAACCCGCTTAAACCCGCTTAAACCCGCTTAAACCCGCTTAAACCCG 840
 DB 781 GAGATAATAAGCATCTTAAACCCGCTTAAACCCGCTTAAACCCGCTTAAACCCGCTTAAACCCG 840
 QY 841 CATATTCAAAGGCAAAAGGTTGCTGCGCATCGGCTGCGCATCGGCTGCGCATCGGCTGCGCAT 900
 DB 841 CATATTCAAAGGCAAAAGGTTGCTGCGCATCGGCTGCGCATCGGCTGCGCATCGGCTGCGCAT 900
 QY 901 TCATACAAACCAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 960
 DB 901 TCATACAAACCAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATAATAATA 960
 QY 961 CAAATCAATTCGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 DB 961 CAAATCAATTCGCGCGGCTGATGATGATGATGATGATGATGATGATGATGATGATGATG 1020
 QY 1021 TGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1080
 DB 1021 TGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTGCTGTG 1080
 QY 1081 CACAGGCGAG 1140
 DB 1081 CACAGGCGAG 1140
 QY 1141 TGATAAATACAGCGGAG 1200
 DB 1141 TGATAAATACAGCGGAG 1200
 QY 1201 ACAAAATACAG 1260
 DB 1201 ACAAAATACAG 1260
 QY 1261 ACCGAG 1320
 DB 1261 ACCGAG 1320
 QY 1321 CTCGCGCT 1380
 DB 1321 CTCGCGCT 1380
 QY 1381 CGAGAGCGGCT 1440
 DB 1381 CGAGAGCGGCT 1440
 QY 1441 ACCTGTSCCAAGAGATGATGAGCGGCTTAAATGCAAAACCAATGATCCGAGAGAGAGAGAG 1500
 DB 1441 ACCTGTSCCAAGAGATGATGAGCGGCTTAAATGCAAAACCAATGATCCGAGAGAGAGAGAG 1500
 QY 1501 CGAACGGGACAGCGGAG 1560

QY 3721 GAACGGCAATCACTCGGCCAGGACTCGATGGAGCCCTGGGAGAAATACATCCGAGAACGG 3780
 DB 3721 GAACGGCAATCACTCGGCCAGGACTCGATGGAGCCCTGGGAGAAATACATCCGAGAACGG 3780
 QY 3781 ACCGAAGCCCAACCGTGTGGAGAGCGGACGGATCGCGGATCACCAAGAGATCCTGGCCCA 3840
 DB 3781 ACCGAAGCCCAACCGTGTGGAGAGCGGACGGATCGCGGATCACCAAGAGATCCTGGCCCA 3840
 QY 3841 CTCGCTGTGGGCAACCGGACTACATAGCTCCGAGAGTCTGGAGAGAGTGGGTACAC 3900
 DB 3841 CTCGCTGTGGGCAACCGGACTACATAGCTCCGAGAGTCTGGAGAGAGTGGGTACAC 3900
 QY 3901 GCAGCTGTGGGCAACCGGACTACATAGCTCCGAGAGTCTGGAGAGAGTGGGTACAC 3960
 DB 3901 GCAGCTGTGGGCAACCGGACTACATAGCTCCGAGAGTCTGGAGAGAGTGGGTACAC 3960
 QY 3961 GCGCTTTCTGGCCCAACAGTCCGCTGGAAAGCGCAACAAAGGTCAACAAGTGGGAGAAAC 4020
 DB 3961 GCGCTTTCTGGCCCAACAGTCCGCTGGAAAGCGCAACAAAGGTCAACAAGTGGGAGAAAC 4020
 QY 4021 SCTGCATATTCGCGCCGAGGCGGAGTTCATCCGCGAGGCTACGAGCTTGATAGAGGCT 4080
 DB 4021 SCTGCATATTCGCGCCGAGGCGGAGTTCATCCGCGAGGCTACGAGCTTGATAGAGGCT 4080
 QY 4081 CTGTGCTGCGCTGACAAAGCGGCTGGGCAAGAGCGTGGAGAGGTCAAGAGCAGCACTT 4140
 DB 4081 CTGTGCTGCGCTGACAAAGCGGCTGGGCAAGAGCGTGGAGAGGTCAAGAGCAGCACTT 4140
 QY 4141 CTTCAAGGGCATCGACTTTCCGAGATGCGGAAGAGCGGCTACATACCGGAAAT 4200
 DB 4141 CTTCAAGGGCATCGACTTTCCGAGATGCGGAAGAGCGGCTACATACCGGAAAT 4200
 QY 4201 CAAGCACCACCGGACACATCCAACTTTGATCCGCTGGATCCGAGAGCTCGGCTCGAA 4260
 DB 4201 CAAGCACCACCGGACACATCCAACTTTGATCCGCTGGATCCGAGAGCTCGGCTCGAA 4260
 QY 4261 TGACTCCACCATGAGCAGCGGCGATGATGCGACCAAGATGACCGGCTTCCACGGCTT 4320
 DB 4261 TGACTCCACCATGAGCAGCGGCGATGATGCGACCAAGATGACCGGCTTCCACGGCTT 4320
 QY 4321 TTTGCAATTTACCTTCGCTGCTTCCTTCGACGACAGCAGCGGCGGATATGACGAGCA 4380
 DB 4321 TTTGCAATTTACCTTCGCTGCTTCCTTCGACGACAGCAGCGGCGGATATGACGAGCA 4380
 QY 4381 TCAGGCGCGGTTTACGCTGGAATGATGCTCTCCATGTCGACCAACACACCGCCG 4440
 DB 4381 TCAGGCGCGGTTTACGCTGGAATGATGCTCTCCATGTCGACCAACACACCGCCG 4440
 QY 4441 CCCCAGATCATTTAGTCAAAATAGTCAACAAAGGGGATAGAAACCAATTGAGTGGGCTT 4500
 DB 4441 CCCCAGATCATTTAGTCAAAATAGTCAACAAAGGGGATAGAAACCAATTGAGTGGGCTT 4500
 QY 4501 GCATTTAAAGGAGCGTGGCTATAGAACTATCATATACATATATATATATATATATAT 4560
 DB 4501 GCATTTAAAGGAGCGTGGCTATAGAACTATCATATACATATATATATATATATATAT 4560
 QY 4561 GGAGCAGTAGAGCGGCGGAGTACGTATATACATACAAATAATATACATATATATATAT 4620
 DB 4561 GGAGCAGTAGAGCGGCGGAGTACGTATATACATACAAATAATATACATATATATATAT 4620
 QY 4621 AT 4680
 DB 4621 AT 4680
 QY 4681 AGATGAACGAGAGGAGCGGAGTACGAGCTTCGACCTTTAAGTGAACATAGTATATATCTT 4740
 DB 4681 AGATGAACGAGAGGAGCGGAGTACGAGCTTCGACCTTTAAGTGAACATAGTATATCTT 4740
 QY 4741 GTGCACCTACTCTCCACAAATAATATATATATATATATATATATATATATATATATAT 4800
 DB 4741 GTGCACCTACTCTCCACAAATAATATATATATATATATATATATATATATATATATAT 4800

QY 4801 ACTGGAATCGAACCTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 DB 4801 ACTGGAATCGAACCTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAG 4860
 QY 4861 TTAACATAAAGAGAGCGGAAATTTACCAACACCACTTCACTCTCTCTCTCTCTCTCTCTCT 4920
 DB 4861 TTAACATAAAGAGAGCGGAAATTTACCAACACCACTTCACTCTCTCTCTCTCTCTCTCTCT 4920
 QY 4921 GATCGGTGGCCGGATTCGAATTCAGACGCTGGTGTGATCCGGGCAATCCCAATTTACCTTCC 4980
 DB 4921 GATCGGTGGCCGGATTCGAATTCAGACGCTGGTGTGATCCGGGCAATCCCAATTTACCTTCC 4980
 QY 4981 CATTCAGAAATTCAGAAATTCGAGGTGCGAGTGGAGAGCAAGCGAGAGCAAAAGTCCGAC 5040
 DB 4981 CATTCAGAAATTCAGAAATTCGAGGTGCGAGTGGAGAGCAAGCGAGAGCAAAAGTCCGAC 5040
 QY 5041 GGCAGCATATAAGCGGCTTTATAAGCCTAATCTAAATCTAAATCTAAATCTAAATCTAAATCT 5100
 DB 5041 GGCAGCATATAAGCGGCTTTATAAGCCTAATCTAAATCTAAATCTAAATCTAAATCTAAATCT 5100
 QY 5101 YATGTATGTCCTGCTATCCAAATTCGCTCTATCAGCTCTCTCACTCTCTCTCTCTCTCTCTCT 5160
 DB 5101 YATGTATGTCCTGCTATCCAAATTCGCTCTATCAGCTCTCTCACTCTCTCTCTCTCTCTCTCT 5160
 QY 5161 CCCCCCTCCCAATCCAAATTCGCTCTATCAGCTCTCTCACTCTCTCTCTCTCTCTCTCTCTCT 5220
 DB 5161 CCCCCCTCCCAATCCAAATTCGCTCTATCAGCTCTCTCACTCTCTCTCTCTCTCTCTCTCTCT 5220
 QY 5221 GTTAGCAACACTTTGTAATGCCAAATGAAATTTGTTAGCCCCACGAGGAAACCGCGGG 5280
 DB 5221 GTTAGCAACACTTTGTAATGCCAAATGAAATTTGTTAGCCCCACGAGGAAACCGCGGG 5280
 QY 5281 GAAATTCACACTTATCTCTGATAGCAACCGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
 DB 5281 GAAATTCACACTTATCTCTGATAGCAACCGGAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 5340
 QY 5341 AAACAGTACGAGAAATTTGTAATCTCTCTGATAGCAACCGGAAAGAAAGAAAGAAAGAAAGAAAG 5400
 DB 5341 AAACAGTACGAGAAATTTGTAATCTCTCTGATAGCAACCGGAAAGAAAGAAAGAAAGAAAGAAAG 5400
 QY 5401 AATCTATCTAGAGTGTGTAGCGCCCTAAGATGTTTTAGTTTATAGACCGTAACCG 5460
 DB 5401 AATCTATCTAGAGTGTGTAGCGCCCTAAGATGTTTTAGTTTATAGACCGTAACCG 5460
 QY 5461 TAATCTAGTTTAAATCTTAACACTAAGCGAGGTACAGTACATTTGTTTGTGTC 5520
 DB 5461 TAATCTAGTTTAAATCTTAACACTAAGCGAGGTACAGTACATTTGTTTGTGTC 5520
 QY 5521 CTAGGTTCTGTTGGAATGCTTAACGGGAAACGATTTGTTTCTCTTTAATAGCTTCA 5580
 DB 5521 CTAGGTTCTGTTGGAATGCTTAACGGGAAACGATTTGTTTCTCTTTAATAGCTTCA 5580
 QY 5581 GTTTGTATGTCGCTGTTTTTATATGACTTATATATAGTCCATCTGAATATTCGTGGA 5640
 DB 5581 GTTTGTATGTCGCTGTTTTTATATGACTTATATATAGTCCATCTGAATATTCGTGGA 5640
 QY 5641 TGGAGCCTATTTTAAATGTGAGATCGAGCTAATTTGAAGGAATACAAACAACTCTGTGT 5700
 DB 5641 TGGAGCCTATTTTAAATGTGAGATCGAGCTAATTTGAAGGAATACAAACAACTCTGTGT 5700
 QY 5701 GCCTTGGCCCAATTTAGTTTAC 5720
 DB 5701 GCCTTGGCCCAATTTAGTTTAC 5720

RESULT 2
 AAT42117
 ID AAT42117 standard; cDNA; 5720 BP.
 AC AAT42117;
 XX
 XX
 DT 22-JAN-1997 (first entry)
 XX


```

DE XX Lats gene encoding large tumour suppressor.
KW XX Lats gene; large tumour suppressor; Drosophila melanogaster;
KW KW fruitfly; polyadenylation site; protein-serine/threonine-kinase;
KW KW cell proliferation; antisense; dominant-negative; cancer;
KW KW degenerative disorder; trauma; growth deficiency; therapy;
KW KW anticumour; vulnary; diagnostic; transgenic plant;
KW KW transgenic animal; growth; senescence; ds.
XX OS Drosophila melanogaster.
XX XX
XX XX
FH FH Key Location/Qualifiers
FT FT misc_feature 1..191
FT FT /*tag= a
FT FT /note= "Sequence from clone cDNA-9"
FT FT 5'UTR 1..1102
FT FT /*tag= b
FT FT 81
FT FT /*tag= c
FT FT /note= "A in genomic sequence"
FT FT 192..5720
FT FT /*tag= d
FT FT /note= "Sequence from clone cDNA-A2"
FT FT misc_feature 302..303
FT FT /*tag= e
FT FT /note= "Intron-1 splice site"
FT FT 330
FT FT /*tag= f
FT FT /note= "T in genomic sequence"
FT FT 606..608
FT FT /*tag= g
FT FT /note= "GGA in genomic sequence"
FT FT 821
FT FT /*tag= h
FT FT /note= "G in genomic sequence"
FT FT 834
FT FT /*tag= i
FT FT /note= "A in genomic sequence"
FT FT 1103..4402
FT FT /*tag= j
FT FT /product= Large tumour suppressor (lats protein)
FT FT misc_feature 1225..1226
FT FT /*tag= k
FT FT /note= "Intron-2 splice site"
FT FT 1371..1818
FT FT /*tag= l
FT FT /note= "lats-al deletion"
FT FT 1447
FT FT /*tag= m
FT FT /note= "C in genomic sequence"
FT FT 1719
FT FT /*tag= n
FT FT /note= "G in genomic sequence"
FT FT 1939
FT FT /*tag= o
FT FT /note= "T in genomic sequence"
FT FT 2137
FT FT /*tag= p
FT FT /note= "C in genomic sequence"
FT FT 2431
FT FT /*tag= q
FT FT /note= "C in genomic sequence"
FT FT 2458
FT FT /*tag= r
FT FT /note= "T in genomic sequence"
FT FT 2551
FT FT /*tag= s
FT FT /note= "G in genomic sequence"
FT FT 2554
FT FT /*tag= t
FT FT /note= "G in genomic sequence"
FT FT 2555
FT FT /*tag= u
FT FT
FT FT /note= "AAC in genomic sequence"
FT FT 2556
FT FT /*tag= v
FT FT /note= "A in genomic sequence"
FT FT 2593
FT FT /*tag= w
FT FT /note= "G in genomic sequence"
FT FT 2646
FT FT /*tag= x
FT FT /note= "GGAGCGGATCAAC insertion in genomic DNA"
FT FT 3488..3489
FT FT /*tag= y
FT FT /note= "Intron-3 splice site"
FT FT 3609..3610
FT FT /*tag= z
FT FT /note= "Intron-4 splice site"
FT FT 3725..3726
FT FT /*tag= aa
FT FT /note= "Intron-5 splice site"
FT FT 3804..3805
FT FT /*tag= ab
FT FT /note= "Intron-6 splice site"
FT FT 3889
FT FT /*tag= ac
FT FT /note= "A in genomic sequence"
FT FT 3937
FT FT /*tag= ad
FT FT /note= "C in genomic sequence"
FT FT 3940
FT FT /*tag= ae
FT FT /note= "T in genomic sequence"
FT FT 4000..4001
FT FT /*tag= af
FT FT /note= "Intron-7 splice site"
FT FT 4021
FT FT /*tag= ag
FT FT /note= "T in genomic sequence"
FT FT 4021
FT FT /*tag= ah
FT FT /note= "C in genomic sequence"
FT FT 4210
FT FT /*tag= ai
FT FT /note= "G in genomic sequence"
FT FT 4300
FT FT /*tag= aj
FT FT /note= "C in genomic sequence"
FT FT 4309
FT FT /*tag= ak
FT FT /note= "C in genomic sequence"
FT FT 4402..5720
FT FT /*tag= al
FT FT 4654..4660
FT FT /*tag= am
FT FT 4797
FT FT /*tag= an
FT FT /note= "G in genomic sequence"
FT FT 4828..4837
FT FT /*tag= ao
FT FT /note= "Deleted in genomic DNA"
FT FT 4871
FT FT /*tag= ap
FT FT /note= "C in genomic sequence"
FT FT 4974
FT FT /*tag= aq
FT FT /note= "T in genomic sequence"
FT FT 5013..5153
FT FT /*tag= ar
FT FT /note= "Region identical to plc-21 transcript"
FT FT 5101
FT FT /*tag= as
FT FT /note= "C in genomic sequence"
FT FT 5104..5114
FT FT /*tag= at
FT FT

```


Db	1381	CGAGCGGGTCCATATCCGTATCCGGTGTGGGGTGTGGAGTGTGGGTGTGGCGAACGG	1440
Qy	1441	ACGTGTSCAAAGATGATGACGGCCCTAATGCCAAACAAACTGATCCGGAAGCCGAGCAT	1500
Db	1441	ACGTGTCCCAAGATGATGACGGCCCTAATGCCAAACAAACTGATCCGGAAGCCGAGCAT	1500
Qy	1501	CGAACGGGACACGGCGAGCAGTCACTACCTGCGGTGCAGTCCGGCTCTGGACTCCGAGAC	1560
Db	1501	CGAACGGGACACGGCGAGCAGTCACTACCTGCGGTGCAGTCCGGCTCTGGACTCCGAGAC	1560
Qy	1561	CGGTAGCTCCCGATCGGACAGCCCCCATTCGCACACACCCACACGCGGAGCTCGAGAC	1620
Db	1561	CGGTAGCTCCCGATCGGACAGCCCCCATTCGCACACACCCACACGCGGAGCTCGAGAC	1620
Qy	1621	GGTGGGTAAATCCAGGTGGAAATGGTGGATTTTCTCCGTGCACAGCGGTTTTCAGTGAAGT	1680
Db	1621	GGTGGGTAAATCCAGGTGGAAATGGTGGATTTTCTCCGTGCACAGCGGTTTTCAGTGAAGT	1680
Qy	1681	GGTCCACCGGGCGCCGCCACGCAATCCACCGCTSCAGCGCGCCACAGGCCCCACC	1740
Db	1681	GGTCCACCGGGCGCCGCCACGCAATCCACCGCTSCAGCGCGCCACAGGCCCCACC	1740
Qy	1741	GCAGTGCGCCACACGCGAGCGGTAGCTGAAGCGGCGATCACGGGCCCTGAAACACCG	1800
Db	1741	GCAGTGCGCCACACGCGAGCGGTAGCTGAAGCGGCGATCACGGGCCCTGAAACACCG	1800
Qy	1801	CCCGCGGGGATAGCGCCACCCACATCAGCGMGGAACATCACTGTATATAACCCCAACACGG	1860
Db	1801	CCCGCGGGGATAGCGCCACCCACATCAGCGMGGAACATCACTGTATATAACCCCAACACGG	1860
Qy	1861	GCTGAAGAACCCGACAGCAGTTGACGACAGCTGAAGTCCCTGAACCTATACCCAGG	1920
Db	1861	GCTGAAGAACCCGACAGCAGTTGACGACAGCTGAAGTCCCTGAACCTATACCCAGG	1920
Qy	1921	CGGAGCAGTGGAGCAGTGTGTGGAGCCACCGCGCCCTACCTAATTCAAGCGGAGCCGG	1980
Db	1921	CGGAGCAGTGGAGCAGTGTGTGGAGCCACCGCGCCCTACCTAATTCAAGCGGAGCCGG	1980
Qy	1981	AGGAGCAGCACCGCGCGCCACACCCAGTTTACAGCGCTTCAATGAGTCTCGGGCAGTC	2040
Db	1981	AGGAGCAGCACCGCGCGCCACACCCAGTTTACAGCGCTTCAATGAGTCTCGGGCAGTC	2040
Qy	2041	GCCACACAATCCCAAAATCGGACTACAGGAAATCCCGGAGCAGTGGGATATCTCGGC	2100
Db	2041	GCCACACAATCCCAAAATCGGACTACAGGAAATCCCGGAGCAGTGGGATATCTCGGC	2100
Qy	2101	CACCTCGCGGGTCCCGAGCCCCATTAATGTGTCTGTCTGCGCGCGCGCGCTGGCGAA	2160
Db	2101	CACCTCGCGGGTCCCGAGCCCCATTAATGTGTGTCTGTCTGCGCGCGCGCGCTGGCGAA	2160
Qy	2161	GCCACAACACAGTGCTACAGCGCCAGGAGTCAGCAGCGCATCATCATGAGAGTGTGAA	2220
Db	2161	GCCACAACACAGTGCTACAGCGCCAGGAGTCAGCAGCGCATCATCATGAGAGTGTGAA	2220
Qy	2221	GAGCAGCAGGTCCRAAAGCCGTGCTGCAACACAGCAGTGGCGGCGCAATCGCCATCGAG	2280
Db	2221	GAGCAGCAGGTCCRAAAGCCGTGCTGCAACACAGCAGTGGCGGCGCAATCGCCATCGAG	2280
Qy	2281	TGCGCTCGGCCAGGAATTCACAGTCCAGTGTGTGGCGCTCCACCTCTTACCCCTCAGAA	2340
Db	2281	TGCGCTCGGCCAGGAATTCACAGTCCAGTGTGTGGCGCTCCACCTCTTACCCCTCAGAA	2340
Qy	2341	GTCCGCGGAGTGGTGCAGCAGCAACACAGCAGCAGCGCGGCCACACAGCAGCAGCA	2400
Db	2341	GTCCGCGGAGTGGTGCAGCAGCAACACAGCAGCAGCGCGGCCACACAGCAGCAGCA	2400
Qy	2401	TCAGCACCAGCAATCCAAACACCACCAACCGCMACCACACCGCCCTTGGTGGGTCTGAAYAG	2460
Db	2401	TCAGCACCAGCAATCCAAACACCACCAACCGCMACCACACCGCCCTTGGTGGGTCTGAAYAG	2460
Qy	2461	CAAGCCCAATTGCTGGAGCCACCGTCTTATGCCAAGAGCATGCAGGCCAAGCGGCCAC	2520
Db	2461	CAAGCCCAATTGCTGGAGCCACCGTCTTATGCCAAGAGCATGCAGGCCAAGCGGCCAC	2520

[illegible]

QY 3601 CATTACAGAGACATCAAGCCTGACAACTACTCATGATGGGACGACACATAAAGCT 3660
DB 3601 CATTACAGAGACATCAAGCCTGACAACTACTCATGATGGGACGACACATAAAGCT 3660
QY 3661 CACCGACTTTGGCCTGTGACCGGATTCGGATGGACGACAACTCGAAGTACTACCGAGGA 3720
DB 3661 CACCGACTTTGGCCTGTGACCGGATTCGGATGGACGACAACTCGAAGTACTACCGAGGA 3720
QY 3721 GAACGGCAATCACTCGCGCCAGGACTCGATGAGAGCCCTGGAGGAATACTCCGAGAACGG 3780
DB 3721 GAACGGCAATCACTCGCGCCAGGACTCGATGAGAGCCCTGGAGGAATACTCCGAGAACGG 3780
QY 3781 ACCGAGCCACCGCTGTGGAGAGCGGACGATCGCGATCACCAAGAGTCTCGGCCCA 3840
DB 3781 ACCGAGCCACCGCTGTGGAGAGCGGACGATCGCGATCACCAAGAGTCTCGGCCCA 3840
QY 3841 CTCGCTGTGGGACCCCGCACTACATAGTCTCCGAGGTGCTGGAGAGTGGGTACAC 3900
DB 3841 CTCGCTGTGGGACCCCGCACTACATAGTCTCCGAGGTGCTGGAGAGTGGGTACAC 3900
QY 3901 GCAGCTGTGGACTACTGGAGCGTGGCGTCACTCTTATGAGATGCTGGTGGGTACGCC 3960
DB 3901 GCAGCTGTGGACTACTGGAGCGTGGCGTCACTCTTATGAGATGCTGGTGGGTACGCC 3960
QY 3961 GCCCTTTCTGCGCAACAGTCCGCTGAAAGCAACAAAGGTATCACTGAGGAGAAAC 4020
DB 3961 GCCCTTTCTGCGCAACAGTCCGCTGAAAGCAACAAAGGTATCACTGAGGAGAAAC 4020
QY 4021 SCTGCAATTTCCGCGCAGCGCGAGTTATCCCGGAGCTACGAGCTTGAAGAGCT 4080
DB 4021 SCTGCAATTTCCGCGCAGCGCGAGTTATCCCGGAGCTACGAGCTTGAAGAGCT 4080
QY 4081 CTGTGCTGCGGTGACAGCGCTGGCAAGAGCTGAGAGAGTCAAGAGCCACGACTT 4140
DB 4081 CTGTGCTGCGGTGACAGCGCTGGCAAGAGCTGAGAGAGTCAAGAGCCACGACTT 4140
QY 4141 CTTCAAGGCACTCGACTTTGCGGACATCGGAAGCAAGAGCGCCCTACATACCGGAAAT 4200
DB 4141 CTTCAAGGCACTCGACTTTGCGGACATCGGAAGCAAGAGCGCCCTACATACCGGAAAT 4200
QY 4201 CAAGCACCACCGACACATCCACTTTGATCCGCTGGATCCGAGAGCTGCGTCCGAA 4260
DB 4201 CAAGCACCACCGACACATCCACTTTGATCCGCTGGATCCGAGAGCTGCGTCCGAA 4260
QY 4261 TGACTCCACCATGAGCAGCGCGATGATGCGACAGATGACCGACATTCACCGGCTT 4320
DB 4261 TGACTCCACCATGAGCAGCGCGATGATGCGACAGATGACCGACATTCACCGGCTT 4320
QY 4321 TTTGCAATTTACCTTCCGCTGCTTTCGACGACAGCAGCGCGGATATGACCGAGCA 4380
DB 4321 TTTGCAATTTACCTTCCGCTGCTTTCGACGACAGCAGCGCGGATATGACCGAGCA 4380
QY 4381 TCAGGCGCGGTTTACGCTCTGAATGAGTCTTCATGTGCGCCCAACACACCCCGC 4440
DB 4381 TCAGGCGCGGTTTACGCTCTGAATGAGTCTTCATGTGCGCCCAACACACCCCGC 4440
QY 4441 CCCCAGATCATTTGTTAGTCAAACTGACAAAAGGGGATGAAACCATTCAGTGGGCTT 4500
DB 4441 CCCCAGATCATTTGTTAGTCAAACTGACAAAAGGGGATGAAACCATTCAGTGGGCTT 4500
QY 4501 GCATTGTAAGGAAGCGTGTATAGATGAACACTATCATATATATAAATATATA 4560
DB 4501 GCATTGTAAGGAAGCGTGTATAGATGAACACTATCATATATATAAATATATA 4560
QY 4561 GGAGACAGTAGGGGAGCTAGCTATATACATAAATAATATACATATATTTGATAT 4620
DB 4561 GGAGACAGTAGGGGAGCTAGCTATATACATAAATAATATACATATATTTGATAT 4620
QY 4621 AT 4680
DB 4621 AT 4680

QY 4681 AGATGAACAGAGAGGAGCGAGTCCAGACCTTCGACCTTTAACTGAACATAGTATATCCTT 4740
DB 4681 AGATGAACAGAGAGGAGCGAGTCCAGACCTTCGACCTTTAACTGAACATAGTATATCCTT 4740
QY 4741 GTGCACCTACTACTCCACAACAATATATATTTTAAATTTGTTAGAAATCAAAAGGRRCA 4800
DB 4741 GTGCACCTACTACTCCACAACAATATATATTTTAAATTTGTTAGAAATCAAAAGGRRCA 4800
QY 4801 ACTGGAATTCGAACCTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4860
DB 4801 ACTGGAATTCGAACCTTTCTGTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCAAGCA 4860
QY 4861 TTAAACTTAATAGAGAGCGAATTTACCCCAACCTTCACCTCTCTCTCTCTCTCTCTCT 4920
DB 4861 TTAAACTTAATAGAGAGCGAATTTACCCCAACCTTCACCTCTCTCTCTCTCTCTCTCT 4920
QY 4921 GATCGGTGGCGGATTCGAACCTGAGAGCTGTTGATCCGCCATCCCATTTKACTTCC 4980
DB 4921 GATCGGTGGCGGATTCGAACCTGAGAGCTGTTGATCCGCCATCCCATTTKACTTCC 4980
QY 4981 CATTCAAGATTTGAGATTTGCGAGGTGTCGATGAGAGCAAGCAAGCAAGCAAGCAAGCA 5040
DB 4981 CATTCAAGATTTGAGATTTGCGAGGTGTCGATGAGAGCAAGCAAGCAAGCAAGCAAGCA 5040
QY 5041 GGCAGCATATAAGCGGCTTTATAAGCCTAATCTAAATCTAAATCTAAATCTAAATCTAA 5100
DB 5041 GGCAGCATATAAGCGGCTTTATAAGCCTAATCTAAATCTAAATCTAAATCTAAATCTAA 5100
QY 5101 YATGATGCTGCTATCCCAATTCGCTATCACTGCTCTCAWCTGWTAGACCCGCCAC 5160
DB 5101 YATGATGCTGCTATCCCAATTCGCTATCACTGCTCTCAWCTGWTAGACCCGCCAC 5160
QY 5161 CCCCCCTCCCTCCCAATTCGCTATCACTGCTCTCAWCTGWTAGACCCGCCAC 5220
DB 5161 CCCCCCTCCCTCCCAATTCGCTATCACTGCTCTCAWCTGWTAGACCCGCCAC 5220
QY 5221 GTTAGCAACCTTTGTAATGCAATGAAATGTTAGCCCCAGAGGAAACACGCGGG 5280
DB 5221 GTTAGCAACCTTTGTAATGCAATGAAATGTTAGCCCCAGAGGAAACACGCGGG 5280
QY 5281 GAAATTCACACTTTCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5340
DB 5281 GAAATTCACACTTTCTCTGATAGCAACGGAAGAAAGAAAGAAAGAAAGAAAGAAAGAA 5340
QY 5341 AAACAGTACGAGAAATTTGTAATCTCTTAATGTAATTTGTAAGCAACAGTTAATTTGT 5400
DB 5341 AAACAGTACGAGAAATTTGTAATCTCTTAATGTAATTTGTAAGCAACAGTTAATTTGT 5400
QY 5401 AATCTATGCTAGAGTTGTGTAGCGCCTAAGATGTTTTTATAGTTTATAGCCGCTAACCG 5460
DB 5401 AATCTATGCTAGAGTTGTGTAGCGCCTAAGATGTTTTTATAGTTTATAGCCGCTAACCG 5460
QY 5461 TAATCTAGTTTAAATCTTAACACTAAGCAGAGTACAGTACATTTGTTTTTTTGTTC 5520
DB 5461 TAATCTAGTTTAAATCTTAACACTAAGCAGAGTACAGTACATTTGTTTTTTTGTTC 5520
QY 5521 GTAGGTCCTGGAAGATGCTTAACGGAACGATTTGTTTTCTTTTAAATAGCTTCA 5580
DB 5521 GTAGGTCCTGGAAGATGCTTAACGGAACGATTTGTTTTCTTTTAAATAGCTTCA 5580
QY 5581 GTTGTATGCTGCTGTTTTTATATACATATATATAGTCCATCTGAATATTCTGGA 5640
DB 5581 GTTGTATGCTGCTGTTTTTATATACATATATATAGTCCATCTGAATATTCTGGA 5640
QY 5641 TGGAGCCTTTTAAATGTGAGTTCGAGCTAATTTGAAGGAAATACAAACAACTCTGTGT 5700
DB 5641 TGGAGCCTTTTAAATGTGAGTTCGAGCTAATTTGAAGGAAATACAAACAACTCTGTGT 5700
QY 5701 GCCTTGGCCAAATAGTTTAC 5720
DB 5701 GCCTTGGCCAAATAGTTTAC 5720

QY	2600	CAACAGCAACTGCAGGCTTTGAGGGTCTCCAGGCACAGGCTCAGA-----	2645
Db	3139	CAACAGCAACTGCAGGCTTTGAGGGTCTCCAGGCACAGGCTCAGAGGAGGGGATCAA	3080
QY	2646	GGGAGCGGGATCAACGGGAGCGGGAAACGGGATACGAGAAGCTGCCAACGGAATCCT	2704
Db	3079	CGGGAGCGGGATCAACGGGAGCGGGAACGGGATACGAGAAGCTGCCAACGGAATCCT	3020
QY	2705	GGCCGGCAGATGCTTTCGCCGCCGCCCTATCAGAGCAACAACAACAACAGCAGGATC	2764
Db	3019	GGCCGGCAGATGCTTTCGCCGCCGCCCTATCAGAGCAACAACAACAACAGCAGGATC	2960
QY	2765	AAACCGCGAGCTGCACAACAACAACAATACAGATAAAGCAACAGCAACTCGCGACGACA	2824
Db	2959	AAACCGCGAGCTGCACAACAACAACAATACAGATAAAGCAACAGCAACTCGCGACGACA	2900
QY	2825	CCACCATTCGCCCTGCCAATACATACTCAACTCTCCACACGGGGCGGAATAGTCG	2884
Db	2899	CCACCATTCGCCCTGCCAATACATACTCAACTCTCCACACGGGGCGGAATAGTCG	2840
QY	2885	GGCGCAGCAACAGGATCCACGGCACCCAGCGCTCTCGTGCACAGCTGCACAAGATC	2944
Db	2839	GGCGCAGCAACAGGATCCACGGCACCCAGCGCTCTCGTGCACAGCTGCACAAGATC	2780
QY	2945	AAGCACGCTTCGCCATCCCGGAGCGCAAGATCTCCAAGGAGAAGGAGGAGCGC	3004
Db	2779	AAGCACGCTTCGCCATCCCGGAGCGCAAGATCTCCAAGGAGAAGGAGGAGCGC	2720
QY	3005	AAGGAGTTCCGCATCAGGCACTACTCGCGCAAGCCTTCAAGTTCTTCATGAGCAGCAC	3064
Db	2719	AAGGAGTTCCGCATCAGGCACTACTCGCGCAAGCCTTCAAGTTCTTCATGAGCAGCAC	2660
QY	3065	ATAGAGAAGCTGATCAAGTCGTATGCCAGCGCACGTATCGCAAGAATCAGCTGGAGAAG	3124
Db	2659	ATAGAGAAGCTGATCAAGTCGTATGCCAGCGCACGTATCGCAAGAATCAGCTGGAGAAG	2600
QY	3125	GAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATCGAGAAATGCTCGAAC	3184
Db	2599	GAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATCGAGAAATGCTCGAAC	2540
QY	3185	CAAAAGGAGAGCAACTACATTCGATGAAGCGCGCAAGATGGACAAGACATACTGGTGC	3244
Db	2539	CAAAAGGAGAGCAACTACATTCGATGAAGCGCGCAAGATGGACAAGACATACTGGTGC	2480
QY	3245	AAACTGAAGCCCATTTGAGTGGGTGCATTTGGCAGGTAAACGCTGGTGACCAAAATCGAT	3304
Db	2479	AAACTGAAGCCCATTTGAGTGGGTGCATTTGGCAGGTAAACGCTGGTGACCAAAATCGAT	2420
QY	3305	ACCTCGAACCATTTGTATCGCATGAAAACCCCTGGCGGAAGCGGACGTTCTCAAGCGGAAT	3364
Db	2419	ACCTCGAACCATTTGTATCGCATGAAAACCCCTGGCGGAAGCGGACGTTCTCAAGCGGAAT	2360
QY	3365	CAGGTGGCACAGTGAAGCGCGAGAGGATATCTTCGGGAGCGGATTAAGCTGGGTG	3424
Db	2359	CAGGTGGCACAGTGAAGCGCGAGAGGATATCTTCGGGAGCGGATTAAGCTGGGTG	2300
QY	3425	GTGAAGTTGTACTACAGCTTCAGAGCAAGGATAATCTGTAATTTGTATGGACTACATA	3484
Db	2299	GTGAAGTTGTACTACAGCTTCAGAGCAAGGATAATCTGTAATTTGTATGGACTACATA	2240
QY	3485	C-----	3485
Db	2239	CCAGGTGAGTGGAAACTTAAAGCAGTCCACGATGTATCCGGAGTCCAAATGTCATTGTG	2180
QY	3486	-----CAGGTGGTGATCTGATGTGCGTGCATCAAACTGGGCATTTTCGAGGAGG	3536
Db	2179	GATTTGTCACAGGTGATCTGATGTGCGTGCATCAAACTGGGCATTTTCGAGGAGG	2120
QY	3537	AACTGCCAGATTTACATCGCGGAGGTACCTGCGCGGTGGACAGCGGTTCACAAATGG	3596
Db	2119	AACTGCCAGATTTACATCGCGGAGGTACCTGCGCGGTGGACAGCGGTTCACAAATGG	2060

QY	3597	GCTTCATT-----	3604
Db	2059	GCTTCATTACACAGGTAATTAGCAGTTGCTTCAGTTATCACAGTCACATTTACATATTGTT	2000
QY	3605	-----CACAGACATCAAGCCTCACAACTACTCATCATAGGACGGACACA	3653
Db	1999	ATGGATTATCCACAGAGACATCAAGCTTGACAACATACTCATGATAGGACGGACACA	1940
QY	3654	TAAAGCTCACCGACTTTGGCGCTGTGCACGGGATTCCGATGGACGCACAACTCGAAGTACT	3713
Db	1939	TAAAGCTCACCGACTTTGGCGCTGTGCACGGGATTCCGATGGACGCACAACTCGAAGTACT	1880
QY	3714	ACCAGGAGAAC-----	3724
Db	1879	ACCAGGAGAGCGGTGAGATGGAAACCCAGCTAACGTAGTGTGTTTAAACAGCTTTTAATAATAT	1820
QY	3725	-----GGCAATCACTCGCGCCAGGACTCGATGGAGCCCTCGGAGGAATA	3768
Db	1819	CATTCCCTACTCTGATAGGCAATCACTCGCGCCAGACTCGATGGAGCCCTCGGAGGAATA	1760
QY	3769	CTCCGAGAACGGACCGGAAGCCCAACCGTGTGTGAG-----	3802
Db	1759	CTCCGAGAACGGACCGGAAGCCCAACCGTGTGTGAGAGGTGAGTTTAAATTAAT	1700
QY	3803	-----AGCGGACGATGCGGATCACCAAA	3827
Db	1699	AGGATCGCGCGTTAATAAAGTATGCCGAATCTCGAGCGACGGATGCGGATCACCAAA	1640
QY	3828	GAGTCTGGCCCCACTCGCTGTGTGGGCACCCGAACTACATAGCTCCGAGGTGCTGGAGA	3887
Db	1639	GAGTCTGGCCCCACTCGCTGTGTGGGCACCCGAACTACATAGCTCCGAGGTGCTGGAGA	1580
QY	3888	GRAGTGGGTACAGCAGCTGTGCGACTACTGAGCGTGGGCGTCATCCTTTATGAGATGC	3947
Db	1579	GAGTGGGTACAGCAGCTGTGCGACTACTGAGCGTGGGCGTCATCCTCTATGAGATGC	1520
QY	3948	TGTTGGGTACAGCGCCCTTTCTGGCCAAACAGTTCCTGGAAACCAACAAA-----	3998
Db	1519	TGTTGGGTACAGCGCCCTTTCTGGCCAAACAGTTCCTGGAAACCAACAAAAGGTTTCGCT	1460
QY	3999	-----AGTCA	4004
Db	1459	CTTTGGATTCTTATAAGCCAAAGTGCAATTACAATAGTACTTGCTTTCATTTTACAGGTCA	1400
QY	4005	TCAACTGGGAGAAAACSCTCGATATTCGCGCGAGCGCGAGTTATCCGCGAGGCTACGG	4064
Db	1399	TCAACTGGGAGAAAACCCCTGCATATTCGCGCGAGCGCGAGTTATCCGCGAGGCTACGG	1340
QY	4065	ACTTGATAAGGAGGCTCTGTGGTGGCTGACACAGCGGCTGGGCAAGAGCGTGGACGAGG	4124
Db	1339	ACTTGATAAGGAGGCTCTGTGGTGGCTGACACAGCGGCTGGGCAAGAGCGTGGACGAGG	1280
QY	4125	TCAAGAGCCACGACTTCTTCAAGGGCATCGACTTTGCGGACATGCGGGAACGACAAAAGCC	4184
Db	1279	TCAAGAGCCACGACTTCTTCAAGGGCATCGACTTTGCGGACATGCGGACGACGAAGGCCG	1220
QY	4185	CCTACATACCGGAAATCAAGCACCCACGAGACACATCCAACTTTGATCCCGTGGATCCGG	4244
Db	1219	CCTATATACCGGAAATCAAGCACCCGAGGACACATCCAACTTTGATCCCGTGGATCCGG	1160
QY	4245	AGAAGCTGCGCTCGAATGACTCCACCATGACACGGCGGATGATGTCGACCAGATGACC	4304
Db	1159	AGAAGCTGCGCTCGAATGACTCCACCATGACACGGCGGATGATGTCGACCAGAGAGCCG	1100
QY	4305	GCAYTTPCCACGGCTTTTTCGAAATTAACCTTCGCTCGCTTCTTCGACGACAAAGCGCCG	4364
Db	1099	GCACTTTPCCACGGCTTTTTCGAAATTAACCTTCGCTCGCTTCTTCGACGACAAAGCGCCG	1040
QY	4365	CGGATATGACGACGATCAGGCGCGGTTTACGTCCTCAATGATGCTCTCCATGTGCC	4424
Db	1039	CGGATATGACGACGATCAGGCGCGGTTTACGTCCTGAAATGATGCTCTCCATGTGCC	980
QY	4425	AACACCAACACCCCGCCCGCCGAATCATTTAGTCAAAATAGTCACAAAAGGGGATAGAA	4484

Db 979 AACACCAACACCCGCGCCGGAATCATGTAGTCAAAATAGTACAAAAAGGGATAGAA 920
QY 4485 ACCATTGAGTGGCTTGCTTGAAGGAGCGTGGCTATAGAAATGAACATCTATATA 4544
Db 919 ACCATTGAGTGGCTTGCTTGAAGGAGCGTGGCTATAGAAATGAACATCTATATA 860
QY 4545 CATTATATAAATTATAGGACAGTATAGAGCGGGAGCTAGCTATATACATACAAATAA 4604
Db 859 CATTATATAAATTATAGGACAGTATAGAGCGGGAGCTAGCTATATACATACAAATAA 800
QY 4605 TACATATATTGTAT 4663
Db 799 TACATATATTGTAT 740
QY 4664 AAAACGAGCGGAGTATAGATGAAA-CGAGAGGAGGAGTACAGACCTTCGACCTTTAAC 4722
Db 739 AAAACGAGCGGAGTATAGATGAAAACGAGAGGAGGAGTACAGACCTTCGACCTTTAAC 680
QY 4723 TGAACATAGTATATCTTGTGCACTACTACTCCACAAATATATATATATATATATAT 4782
Db 679 TGAACATAGTATATCTTGTGCACTACTACTCCACAAATATATATATATATATATAT 620
QY 4783 AGAATTCAAAAGGRCGAACCTGAAATCGAACCTTTCTGTGTCTCAAGCAAGCAAGC 4842
Db 619-AGAATTCAAAAGGAGCAACTGGAATCGAACCTTTCTGTGTCT-----CAAAGC 570
QY 4843 AAAGCAAAACAAACGCTTAACTAAAYGAGAGCGGAATTTACCAACACACTTCACTCC 4902
Db 569 AAAGCAAAACAAACGCTTAACTAAACGAGAGCGGAATTTACCAACACACTTCACTCC 510
QY 4903 TCTCTTTTCCACCTCCGATCGGTGGCGGATTCGAACCTCAGCAGGCTGGTTCATCCG 4962
Db 509 TCTCTTTTCCACCTCCGATCGGTGGCGGATTCGAACCTCAGCAGGCTGGTTCATCCG 450
QY 4963 GCATCCCAATTKATCTCCCAATTCAGAAATGAGATGCGAGGTGCGATGGAGAGCAAC 5022
Db 449 GCATCCCAATTKATCTCCCAATTCAGAAATGAGATGCGAGGTGCGATGGAGAGCAAC 390
QY 5023 GGAGACCAAAAGTCGACGCGCATATAGCGGCTTATAGCCCTATCTAAATCTA 5082
Db 389 GGAGACCAAAAGTCGACGCGCATATAGCGGCTTATAGCCCTATCTAAATCTA 330
QY 5083 AACTGGGAGAACAGGACCAATGTA-----TGCTCTGCTCCTCAATTCGCTATC 5131
Db 329 AACTGGGAGAACAGGACCAATGTAATGTAATGTAATGTAATGTAATGTAATGTAATG 270
QY 5132 ACTGCTCTCAWCTGWTAGCAGCCGCCACCCGCC-----CTCCCATCCAAAGAACAA 5187
Db 269 ACTGCTCTCAWCTGWTAGCAGCCGCCACCCGCC-----CTCCCATCCAAAGAACAA 210
QY 5188 ACTTAGACCTAGCTATGTGCAAAAGCTAGCAATGTTAGACCAACTTGTGAATGCCAAAT 5247
Db 209 ACTTAGACCTAGCTATGTGCAAAAGCTAGCAATGTTAGACCAACTTGTGAATGCCAAAT 150
QY 5248 GAAATGTTTATAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 5307
Db 149 GAAATGTTTATAGCCCAAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 90
QY 5308 AAACGGAAG 5367
Db 89 AAACGGAAG 35
QY 5368 TTAATGTAATATTGTAAGACACGTTAATTGTA 5401
Db 34 TTAATGTAATATTGTAAGACACGTTAATTGTA 1

RESULT 4
ID ABL03169
ABL03169 standard; cDNA; 3319 BP.
XX
AC ABL03169;

XX 26-MAR-2002 (first entry)
XX Drosophila melanogaster expressed polynucleotide SEQ ID NO 3989.
DE Drosophila; developmental biology; cell signalling; insecticide;
KW pharmaceutical; gene; ss.
XX Drosophila melanogaster.
XX WO200171042-A2.
XX 27-SEP-2001.
XX 23-MAR-2001; 2001WO-US09231.
XX 23-MAR-2000; 2000US-191637P.
PR 11-JUL-2000; 2000US-0614150.
XX (PEKE) PE CORP. NY.
PA Venter JC, Adams M, Li PWD, Myers EW;
PI WPI; 2001-656860/75.
DR P-PSDB; ABB59066.
XX New isolated nucleic acid detection reagent for detecting 1000 or more genes from Drosophila and for elucidating cell signalling and cell-cell interactions.
XX Claim 1; SEQ ID NO 3989; 21pp + Sequence Listing; English.
XX The invention relates to an isolated nucleic acid detection reagent capable of detecting 1000 or more genes from Drosophila. The invention is useful in developmental biology and in elucidating cell signalling and cell-cell interactions in higher eukaryotes for the development of insecticides, therapeutics and pharmaceutical drugs. The invention discloses genomic DNA sequences (ABL16176-ABL30511), expressed DNA sequences (ABL01840-ABL16175) and the encoded proteins (ABB57737-ABB72072).
XX The sequence data for this patent did not form part of the printed specification, but was obtained in electronic format directly from WIPO at ftp.wipo.int/pub/published_pct_sequences.
XX Sequence 3319 BP; 860 A; 1032 C; 933 G; 494 T; 0 other;
SQ
Query Match 56.7%; Score 3242.2; DB 23; Length 3319;
Best Local Similarity 98.7%; Pred. No. 0;
Matches 3276; Conservative 16; Mismatches 9; Indels 18; Gaps 2;
QY 1103 ATGCATCCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1162
Db 1 ATGCATCCAGCGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 60
QY 1163 CTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGTACAAAATACCATAGGAATAAT 1222
Db 61 CTCGAGAGCATCAAGCAGGACCTAACCCGATTTGAAGTACAAAATACCATAGGAATAAT 120
QY 1223 CAGAATTACACCTCTGCGATACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 1282
Db 121 CAGAATTACACCTCTGCGATACAGGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGG 180
QY 1283 GACTATCACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 1342
Db 181 GACTATCACACCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 240
QY 1343 GTGGTATACCGTA 1402
Db 241 GTGGTATACCGTA 300
QY 1403 TCCGGTGTGGCGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 1462
Db 301 TCCGGTGTGGCGGTGGAGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 360

Db 2521 CACAGAGACATCAAGCTGCAACATATCTATCATGATAGGACGACACATAAAGCTCACC 2580
 QY 3665 GACTTTGGCTGTGCACGGATTCCGATGGACGACACAACTGGAAGTACTACACGAGAAC 3724
 Db 2581 GACTTTGGCTGTGCACGGATTCCGATGGACGACACAACTGGAAGTACTACACGAGAAC 2640
 QY 3725 GGCAATCACTCGGCCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGGACCG 3784
 Db 2641 GGCAATCACTCGGCCAGGACTCGATGGAGCCCTGGAGGAATACTCCGAGAACGGACCG 2700
 QY 3785 AAGCCACCGTGTGGAGAGCGGAGGATGGCGGATACCAAGAGTCTTGCCCACTCG 3844
 Db 2701 AAGCCACCGTGTGGAGAGCGGAGGATGGCGGATACCAAGAGTCTTGCCCACTCG 2760
 QY 3845 CTGGTGGGACCCCGAATACATAGCTCCGAGGTCTGGAGAGAGTGGGTACACCCAG 3904
 Db 2761 CTGGTGGGACCCCGAATACATAGCTCCGAGGTCTGGAGAGAGTGGGTACACCCAG 2820
 QY 3905 CTGTGCGACTACTGGAGCGTGGGCTCATCTTATYAGATGCTGTGGGTTCAGCCGCC 3964
 Db 2821 CTGTGCGACTACTGGAGCGTGGGCTCATCTCTATGAGATGCTGTGGGTTCAGCCGCC 2880
 QY 3965 TTTCTGCGCAACAGCTCGCTGGAAACGCAACAAAGTCTCAACTGGGAGAAACCTG 4024
 Db 2881 TTTCTGCGCAACAGCTCGCTGGAAACGCAACAAAGTCTCAACTGGGAGAAACCTG 2940
 QY 4025 CATATTCGCGGACGAGCTGATATCCGAGGCTACGACTGTGATAGGAGGCTCTGT 4084
 Db 2941 CATATTCGCGGACGAGCTGATATCCGAGGCTACGACTGTGATAGGAGGCTCTGT 3000
 QY 4085 GCCTGCGCTGACAGCGCTGGGCAAGAGCGTGGACAGGTCAAGAGCCACGACTTCTTC 4144
 Db 3001 GCCTGCGCTGACAGCGCTGGGCAAGAGCGTGGACAGGTCAAGAGCCACGACTTCTTC 3060
 QY 4145 AAGGCATGACTTTGGGACATGCGGAGAGCAAGACGCGCTTACATACCGGAAATCAAG 4204
 Db 3061 AAGGCATGACTTTGGGACATGCGGAGAGCAAGACGCGCTTACATACCGGAAATCAAG 3120
 QY 4205 CACCCACGACACATCAACTTTGATCCGCTGATCCGAGAGCTCGCTCAATGAC 4264
 Db 3121 CACCCACGACACATCAACTTTGATCCGCTGATCCGAGAGCTCGCTCAATGAC 3180
 QY 4265 TCACCATGACAGCGCGGATGATGTCAGCAGAAATGACCGCCTTCCACCGGCTTTTC 4324
 Db 3181 TCACCATGACAGCGCGGATGATGTCAGCAGAAATGACCGCCTTCCACCGGCTTTTC 3240
 QY 4325 GAATTTACCTTCGCTGCTTCTTCGAGCAAGACGCGCGGATATGACGGAGCATGAC 4384
 Db 3241 GAATTTACCTTCGCTGCTTCTTCGAGCAAGACGCGCGGATATGACGGAGCATGAC 3300
 QY 4385 GCGCGGTTTACGCTCTGAA 4403
 Db 3301 GCGCGGTTTACGCTCTGAA 3319

RESULT 5
 AAT42120
 ID AAT42120 standard; cDNA; 3155 BP.
 AC AAT42120;
 XX
 DT 31-JAN-1997 (first entry)
 XX M-lats2 gene encoding large tumour suppressor.
 DE Mouse; m-lats2 gene; large tumour suppressor; fetal brain;
 XX protein-serine/threonine-kinase; cell proliferation; antisense;
 KW dominant-negative; cancer; degenerative disorder; trauma;
 KW growth deficiency; therapy; antitumour; vulnerability; diagnostic;
 KW transgenic plant; transgenic animal; growth; senescence; ds.
 XX Mus musculus.
 OS

XX Key Location/Qualifiers
 FH 1..2943
 FT /*tag- a
 FT /product= m-lats2 protein
 XX
 PN WO9630402-A1.
 PD 03-OCT-1996.
 PF 26-MAR-1996; 96WO-US04101.
 PR 27-MAR-1995; 95US-0411111.
 XX (UYVA) UNIV YALE.
 XX Tao W, Wang W, Xu T, Yu W, Zhang S;
 XX WPI; 1996-455275/45.
 DR P-PSDB; AAW06287.
 XX New isolated large tumour suppressor gene - used to develop prods.
 PT for inhibiting cell proliferation or for enhancing proliferation
 XX Claim 8; Page 133-137; 215pp; English.
 XX This sequence encodes a mouse large tumour suppressor m-lats2
 CC protein, and has been isolated from a newborn mouse brain phase
 CC lambda-ZAP cDNA library using a 2.2-kb DNA probe from the Drosophila
 CC lats gene (AAT42117). A homologous mouse sequence has also been
 CC isolated (m-lats, AAT42119). The gene encodes a putative protein-
 CC serine/threonine-kinase, and inhibits cell proliferation and plays a
 CC crucial role throughout development. Activators or inhibitors of
 CC lats function (e.g. an antisense oligonucleotide or dominant-negative
 CC lats fragment) may be used in therapy of cancer or other
 CC proliferative disorders, degenerative disorders, trauma, growth
 CC deficiency, etc., and fragments of the gene may be used as diagnostic
 CC probes. A lats-inhibitor sequence may be expressed in a transgenic
 CC plant or farm animal to confer increased growth and inhibit
 CC senescence.
 XX Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;

Query Match 9.9%; Score 567.8; DB 17; Length 3155;
 Best Local Similarity 64.3%; Pred No 6.3e-110;
 Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;
 QY 2869 GGGCGGAATAGTTCGGGCGGACGACGATCCCGGACCGACCGCTCTCTCGTGCAC 2928
 Db 1329 GGGCACTGTACTAGACGGGAGTGACAAAGAGCCAAAGGTCGGAAGGAGACAAAGCTGG 1388
 QY 2929 CAGCTGCAAGAAGATCAAGCAGCTCGCCATCCCGGAGGCGGAGAGATCTCCAAGGA 2988
 Db 1389 CAGACACAAAAGCAGATTCAGACCTCCCGGTGCTTCGCAAGAAATAGC----- 1440
 QY 2989 GAAGGAGGAGGAGCGCAAGGAGTTCCGCATCAGGACGACTCTGCGCGCAAGCTTCAAGTT 3048
 Db 1441 -AGAGATGAAGAGAGAGAGAGTCTCGCATCAAGAGTACTCCCTTATGCTTCAATT 1499
 QY 3049 CTTTCATGGAGCAGCATATAGAACGATGATCAAGTCTGATCGCAGCGCAGTATCGCAA 3108
 Db 1500 CTTTCATGGAGCAACACGTCGGAATGTATCAAAACCTTACGACGAGAGGTCAGCCCGAG 1559
 QY 3109 GAATCAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCCGATCAGACCCCAATCGAGAT 3168
 Db 1560 GCTACAGCTGGAGCAGGAATGCGCAAGCTGGGCTCTGTGAGGCGCAGCAGGACGAGAT 1619
 QY 3169 GAGCAAAATGCTGAACCAAGAGGAGACAACTACATTCGATTTGAAGCGCGCAAGATGGA 3228
 Db 1620 GAGAGATCTCTTACCAAGAGGAGTCTAATCAACCGGCTGAAGAGGCGCAAGATGGA 1679
 QY 3229 CAAGAGCATGTTGTTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTAAAGCT 3288
 Db 3229 CAAGAGCATGTTGTTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTAAAGCT 3288

Db 1680 CAAGTCACCTGTTGTGAAATCAAGACTCTAGGCATCGGTGCCTTTGGGAAAGTGTGCCT 1739
QY 3289 GGTGAGCAAAATCGATACCTCGAACCATTTGTATGGATGATAAAGCCCTGCGGAAAGCGGA 3348
Db 1740 CGCTTGTGAAGCTGGA---CACTCAGCGTCTGTACGCCATGAAGACTCTCAGGAAGGA 1796
QY 3349 CGTCTCAAGCGGAATCAGTGTGCGACACGCTGAAGCGGAGGAGGATATCTCCGCGAAGC 3408
Db 1797 TGTCTTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTGCTGAAGC 1856
QY 3409 CGAATAAATCTGGTGGTGAAGTGTACTACAGTCTCCAGGACAAGGATATCTGTACTT 3468
Db 1857 AGACATGAGTGGTGGTGAAGTGTACTACTCTTCCAGGACAAGGAGAGGCTGTACTT 1916
QY 3469 TGTGATGGACTACATACCAAGGTGGTGTATGTGTGCTCTCATCAAACTGGGCAATTT 3528
Db 1917 TGTGATGGACTACATACCAAGGCGGGATATGATGAGCTCTGTATCAGGATGAGTCTT 1976
QY 3529 CGAGGAGAACTGGCCAGATTTACATCGCGAGGTCTACTGGCGGCTGGACAGCGTTCA 3588
Db 1977 CCTGAGCACCTGGCCGCTTCTACATTCAGAGTTGACCCCTGGCCATTGAAAGTGTCCA 2036
QY 3589 CAAATGGGCTTCATTCACAGAGACATCAAGCTGACACATCTCATCGATAGGACGG 3648
Db 2037 CAAGATGGGCTTTATCCAGCGGACATCAAGCTGACACATCTCATCGCTGGATGG 2096
QY 3649 ACATATAAGCTACCGACTTTGGCTGTGACAGGATTCGATGAGCAGCAACACTCGAA 3708
Db 2097 TCATATAAGCTGACAGATTTTGGCTGTGACAGGATTCAGTGGACTCACAATTCGA 2156
QY 3709 GTACTACAGGAGAAAGCTGCTGGCCGAGGACTCGATGGAGCCCTGGAGGAAT- 3767
Db 2157 GTACTACC---AGAAAGGAAACCATATGACAGGACAGCATGGAGCCCGGTGACCTGTG 2213
QY 3768 -----ACTCCGAGAACCGGACCGACCGCTGCTGGAGAGCGGAGAT 3813
Db 2214 GGACGATGTTTCCACTGTGCTGTGGACAGAGTTAAGACCTTGGACAGAGGCGGA 2273
QY 3814 GCGGATACCAAGAGTCTTGGCCGACCTGCTGTGGGACCCCGGAACTACATAGCTCC 3873
Db 2274 GAAGCAGCACAGAGTCTGGCACATTTCTTGTGGGACACCAAAATACATCGCTCC 2333
QY 3874 CGAGTGTCTGGAGAGAGTGTGACAGCAGCTGTGCGGACTACTGGAGCGTGGGCTCAT 3933
Db 2334 GGAGTGTCTTCCGAAAGGATGACAGCAGCTGTGTACTGGTGGAGCGTGGTGTGAT 2393
QY 3934 CCTTAYAGATGCTGTGGGTGACGCGCCCTTCTTGGCCAAACAGTCCGCTGGAACGCA 3993
Db 2394 TCTCTTTGAGATGCTGTGGGACCGCCCTTCTTGGCCCCACCCACAGACAGCGCA 2453
QY 3994 ACAAGGTATCATCTGGGAGAAACSTGTATATTCGCGGAGGCGGAGTTATCCGG 4053
Db 2454 GCTGAAGGTGATCAACTGGGAGACAGCTGTATATCTCCTAGCAGGTGAGGCTCAGCGC 2513
QY 4054 CGAGGCTACGAGTGTGATAAGAGAGCTGTGCTGCTGCTGACAGAGCGGTGGGCA---A 4110
Db 2514 TGAGCCGAGACCTCATACAGAGCTGTGTGCGGCTGACTCCGCGCTGGGAGGGA 2573
QY 4111 GAGCGTGGACAGGTCAAGAGCAGCAGCTTCTTAAAGGCGCATCGACTTTC---GGACAT 4167
Db 2574 TGGGCGAGATGACCTCAAGGCACACCCGCTCTTCAACACCATCGACTTTTCCGCTGACAT 2633
QY 4168 GGGAGCAGAAAGCGCCCTCATACCGGAATCAAGCAACCCACCGGACATCCAACTT 4227
Db 2634 CCGAAGCAGGCTGACCCCTAGTCCCGCCACCATCAAGCCACCCATGAGACCTCCAACTT 2693
QY 4228 TGATCCGCTGATCGGAGAGCTGCGCTCGAATGACTCCACCATGACAGCGGCGATGA 4287
Db 2694 TGACCCGCTGATGAAGAAGACCCCTGGCAGAGGCGGAGAGCGGCAAGGCGCTG 2753
QY 4288 TGTCCACAGAAATGACCCACATTTTC-----CACGGCTTTTTCGAATTTACCTT 4335
Db 2754 GGACAGCTGCGCTCCCGCCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGATTCACCTT 2813

QY 4336 CCCTCGCTTCTTCGACGACAA 4356
Db 2814 CCGCAGGTTCTTCGATGACAA 2834

RESULT 6

AAZ51507
ID AAZ51507 standard; DNA; 3155 BP.
XX
AC AAZ51507;
XX
DT 21-JUN-2000 (first entry)
XX
Mouse Lats2 (large tumour suppressor) DNA.

Mouse; Lats2; large tumour suppressor; cytostatic; vulnery;
cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
treatment; prevention; screening; cancer; skin; ovarian tumour;
soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
dysplasia; degenerative disorder; growth deficiency; physical trauma;
hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

OS Mus musculus.

XX Key Location/Qualifiers

1..2943

/*tag= a

/product= "Lats2 protein"

970..1920

/*tag= b

/note= "This region is erroneously repeated in the

mouse Lats2 DNA sequence shown in figure 14"

1921..2960

/*tag= c

/note= "This region is missing in the mouse Lats2

DNA sequence shown in figure 14"

WO200010602-A1.

02-MAR-2000.

18-AUG-1999; 99WO-US19068.

18-AUG-1998; 98US-0096996.

18-AUG-1998; 98US-0096997.

(UYUA) UNIV YALE.

Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

Trenchalk GS, Stewart RA;

WPI: 2000-246496/21.

P-PSDB: AAY70392.

Use of lats proteins, complexes of lats and cdc2 for treating cancer

that is refractory to treatment by standard chemotherapy and radiation

therapy, and disorders associated with aberrant levels of cdc2 activity

Claim 44; Pages 112-117; 134pp; English.

The present sequence is a DNA encoding mouse Lats2 (large tumour

suppressor) protein which is a cell overproliferation inhibitor and a

negative regulator of cell cycle-dependent kinase cdc2/cyclin A.

The present sequence is useful for treating cancer that is refractory

to standard chemotherapy or radiation therapy such as hyperplasia,

metaplasia, or dysplasia, and disorders associated with aberrant

levels of cdc2 activity. Conditions treated by promoting cdc2 function

include degenerative disorders, growth deficiencies, hypoproliferative

disorders, physical trauma, lesions, and wounds. An animal model

preferably a mouse, in which a lats gene has been disrupted by homologous

CC recombination, e.g. a lats knock-out mouse, is used for screening
 CC compounds that can be used to treat or prevent cancer, particularly
 CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
 CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
 CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
 XX

Sequence 3155 BP; 751 A; 924 C; 894 G; 586 T; 0 other;
 Query Match 9.9%; Score 567.8; DB 21; Length 3155;
 Best Local Similarity 64.3%; Pred. No. 6.3e-110;
 Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;

QY 2869 GGGCCGGAATAGCTCGGCGGAGCAAGATCCACCGGACACCGCTCTCTCGTCGAC 2928
 DB 1329 GGGCACTACTAGACGGGAGTGAAGAGGCGACAAAGGTCGGAAGGAGACAAAGCTGG 1388
 QY 2929 CAGCTGCAAGAGATCAAGCAGCGCTCGCCCATCCGCGAGCGCAAGAGATCTCAAGGA 2988
 DB 1389 CAGAGCAAAAGCAGATTCAGACCTCCCGGTGCTGTCCGCAAGATAGC----- 1440
 QY 2989 GAAGGAGGAGGAGCGCAAGGAGTTCGGCATCAGGAGTACTCGCGCAAGCCTTCAAGTT 3048
 DB 1441 -AGAGATGAGAGAGAGAGAGTCTCGCATCAAGAGTTACTCCCTTTATGCTTCAAAAT 1499
 QY 3049 CTTATGAGGAGCAGACATAGAGAAGCTGATCAAGTCGTATCCCGCAGCGACGATATCGCAA 3108
 DB 1500 CTTATGAGGAGCAGACATAGAGAAGCTGATCAAGTCGTATCCCGCAGCGACGATATCGCAA 3108
 QY 3109 GAATCAGCTGAGAGGAGATGACAAAGTGGGACTGCGCGATCAGACCGCAAAATCGAGAT 3168
 DB 1560 GCTACAGCTGAGAGGAGATGACAAAGTGGGACTGCGCGATCAGACCGCAAAATCGAGAT 3168
 QY 3169 GAGGAAATGCTGAAACCAAGGAGGAGCAACTACATTCGATTGAGCGCGCAAGATGGA 3228
 DB 1620 GAGGAGATCTCTACAGAGAGAGTCTAATACACCGGCTGAAGAGCGGCAAGATGGA 1679
 QY 3229 CAAGAGCATGTTGCTCAAACTGAAGCCGATGAGAGTGGTGCATTTGGCGAGGTAAACGCT 3288
 DB 1680 CAAGTCCATGTTGTTGAAATCAAGACTTAGGCAATCGGTGCTTTGGGAAAGTGTGCT 1739
 QY 3289 GGTGAGCAAAATCGATACCTCGAACCATTTGTATCGGATGAAACCTTCGGGAAAGCGGA 3348
 DB 1740 CGTTGTAAGCTGGA--CACTCAGGCTGTGACGCCATGAAGACTCTCAGGAAGAAGGA 1796
 QY 3349 CGTTCTCAAGGGAATCAGGTGGACACGCTGAAGCGCGAGAGGGATATCCTCGCGGAAGC 3408
 DB 1797 TGTCTGAACCGGATCAAGTGGCCCATGTCAAGCTCAGAGGACATCTCGGCTGAAGC 1856
 QY 3409 CGACAATACTGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAAGGATATCTGTACTT 3468
 DB 1857 AGACAATAGTGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAAGGATATCTGTACTT 1916
 QY 3469 TGTATGACTACATACAGGTGGTGTATCTGATGCTGCTCATCAACTGGGCGATTTT 3528
 DB 1917 TGTATGACTACATACAGGTGGTGTATCTGATGCTGCTCATCAACTGGGCGATTTT 1976
 QY 3529 CGAGGAGCAACTCGCCGATTTCTACATCGCGAGGTCACTGCGCGGTGGAGAGGTTCA 3588
 DB 1977 CCGTGAGCACTGGCCGCTTCTACATTCGAGAGTTGACCCCTGGCCATGAAAGTGTCA 2036
 QY 3589 CAAATGGGCTTCAATTCACAGAGACATCAAGCCTGACACATACATCATCGATAGGACGG 3648
 DB 2037 CAAATGGGCTTCAATTCACAGAGACATCAAGCCTGACACATACATCATCGATAGGACGG 2096
 QY 3649 ACACATAAGCTCACCAGCTTTGGCTGTGACGGGATTCGATGGAGCGACAACTCGAA 3708
 DB 2097 TCATATTAAAGTGACAGATTTGGCCCTCTGCACTGGATTCAGGTGGACTCAAAATTCGAA 2156
 QY 3709 GTACTTACAGGAGACGCAATCACTCGCGCAGGACTCGATGGAGCGCTGGGAGGAT- 3767
 DB 2157 GTACTACG--AGAAAGGACACCATGAGACAGACAGCATGGAGCCCGGTGACCTCTG 2213
 QY 3768 -----ACTCCGAGAACGGACCGAAGCCACCGTGTCTGGAGAGCGCAGCGAT 3813

DB 2214 GGACGATGTTTCCAACTGTGCGTGTGGAGACAGGTTAAAGACCCCTGGAGCAGGGCGCA 2273
 QY 3814 GGGCCATCACCACAAAGAGTCTGGCCGCTGCTGGTGGGACCCCGAAGTACTATAGTCTCC 3873
 DB 2274 GAAGCAGCACCAGAGAGTGCCTGGCAGATCTCTTCTGGGACACCAAAATTACATCGCTCC 2333
 QY 3874 CCAGCTGCTGGAGAGAGTGGGTACACGAGCTGTGCGACTACTGAGGCTGGGCTCAT 3933
 DB 2334 GGAGGTGCTTCTCCCAAGAGGTACACGAGCTCTGTACTGTGGAGGCTCGGTGTGAT 2393
 QY 3934 CCTTAYGAGATGCTGGGTGAGCGGCTTCTGCGCAACAGTTCGCTGGAAACGCA 3993
 DB 2394 TCTCTTTGAGATGCTGGTGGCAGCGGCTTCTTGGCCCCCACCACCCACAGACGCA 2453
 QY 3994 ACAAAAGTCACTCAACTGGGAGAAAACSTGCATATTCGCGCGCAGCGGATTCCTCCG 4053
 DB 2454 GCTGAAGGTGATCAACTGGGAGAGCAGCTGCATATCTCCCTACGCAAGGTGAGCTCAGCGC 2513
 QY 4054 CGAGCTACGGACTTGAAGGAGGCTGTGCTGCTGCGCTGACAAAGCGCTGGCA--A 4110
 DB 2514 TGAGCCCGGAGACCTCATCAGCAAGCTGTGCTGCGGCTGACTGCGGCTGGGCGAGGA 2573
 QY 4111 GAGCTGAGCAGAGTCAAGAGCCACGACTTCTTCAAGGGCATCGACTTTCG---GGACAT 4167
 DB 2574 TGGGGCAGATGACCTCAAGGCAACCCGCTTCTTCAACACCATCGACTTTTCCCGTGACAT 2633
 QY 4168 GCGGAAGCAGAAAGCGGCTTACATACCGGAAATCAAGCAACCCACGAGGACACATCAACTT 4227
 DB 2634 CGAAAGAGGCTGCAACCTAGTCCCAACCATCAGCCACCCCATGGACACCTCCAATTT 2693
 QY 4228 TGATCCGCTGGATCGGAGAGCTGCGTTCGAATCACTCCACCATGAGCAGCGGATGA 4287
 DB 2694 TGACCCGCTGGATGAAGAAAGCCCTGGCAGCGGAGAGAGCGCCAAAGCGCTG 2753
 QY 4288 TGTCCAGCAGAAATACCCGACATTC-----CACGGCTTTTTCGAATTTACCTT 4335
 DB 2754 GGACACGCTGGCTCCCGCAGCAGCAAGCATCCAGCAGCGCTTCTATGACTTCCACTT 2813
 QY 4336 CGCTGCTTCTTCGAGGACAA 4356
 DB 2814 CGCAGGTTCTTCGATGACAA 2834

RESULT 7
 AAX87397
 ID AAX87397 standard; cdna; 5276 BP.
 XX
 AC AAX87397;
 DT 08-OCT-1999 (first entry)
 XX
 DE Human WART2 CDNA.
 XX
 KW WART2; WART2; WART orthologue; human; signal transduction;
 KW protein kinase; cancer; tumour; diagnosis; therapy; ss..
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT 5'UTR 1..374
 FT /*tag= a
 FT CDS 375..3641
 FT /*tag= b
 FT 3'UTR 3642..5276
 FT /*tag= c
 XX
 PN WO9937787-A2.
 XX
 PD 29-JUL-1999.
 XX
 PF 20-JAN-1999; 99WO-US01145.
 XX

QY 4186 CTACATACCGGAATCAACACCCACACACATCCAACTTTGATCCGTCGATCCGGA 4245
 DB 3350 CTACGTTCCACCATCAGCCACCCACACCTCGAATTCGACCCGTCGATGAAGA 3409
 QY 4246 GAAGTCGCTCGAATGACTCCACCATGAGACGCGCGATGATGTC-----GA 4293
 DB 3410 AAGCCCTTGAACGATGCGCAGGAGGTAGCACCACCAAGGCTCGGACACACTCACCTCGCC 3469
 QY 4294 CCAGATGATGACGACGCTTCCACGCGCTTTTTCGAATTTACTTTCGCTCTTCGACGA 4353
 DB 3470 CAATACAGATCTTGACGACGCAATTTACGAATTCACCTTCCGAGGTTCTTTGATGA 3529
 QY 4354 CAA 4356
 DB 3530 CAA 3532

RESULT 9

AAA59130
 ID AAA59130 standard; DNA; 5486 BP.
 AC AAA59130;
 DT 07-NOV-2000 (first entry)
 DE DNA encoding a tumour suppressor protein hGHITS2.

Human; growth hormone inhibited tumour suppressor protein; hGHITS;
 antineoplastic; dwarfism; gigantism; acromegaly; angiopathy;
 diabetic nephropathy; cardiopathy; tumour; breast cancer;
 renal adenocarcinoma; colorectal cancer; leukaemia; ss.

OS Homo sapiens.

Key Location/Qualifiers
 CDS 387..3653
 FT /*tag= a
 FT /product= "tumour suppressor protein"

EP1022333-A1.

26-JUL-2000.

07-OCT-1999; 99EP-0119199.

25-JAN-1999; 99JP-0016223.

(JCRP-) JCR PHARM CO LTD.

Koga J, Kono K, Zolotaryov FN;

WPI; 2000-516013/47.

P-PSDB; AAB07664.

New human growth hormone inhibited tumour suppressor genes 1 and 2 for
 diagnosis of dwarfism, gigantism, acromegaly, angiopathy, diabetic
 nephropathy or cardiopathy

Claim 1; Page 28-39; 59pp; English.

The present sequence encodes a human growth hormone inhibited tumour
 suppressor (hGHITS) protein. The hGHITS proteins can be used in an
 antineoplastic pharmaceutical preparation. Probes for the hGHITS
 DNA sequences can be used in diagnostic pharmaceutical preparations.
 The diagnostic pharmaceutical preparations can be used for examining
 expression of hGHITS genes in dwarfism, gigantism, acromegaly,
 angiopathy, diabetic nephropathy or cardiopathy, or in malignant
 tumours including breast cancer, renal adenocarcinoma, colorectal
 cancer, and leukaemia. Antibodies against the proteins can be used in
 a diagnostic pharmaceutical preparation for examining expression of
 a tumour suppressor gene.

SO Sequence 5486 BP; 1380 A; 1451 C; 1388 G; 1267 T; 0 other;
 Query Match 9.8%; Score 558.6; DB 21; Length 5486;
 Best Local Similarity 64.1%; Pred. No. 7e-108;
 Matches 964; Conservative 4; Mismatches 493; Indels 42; Gaps 7;
 QY 2884 GGGCGGACGACGATCCAGCCACCCGCTCTCGTCGACACGATCGAAGAAGAT 2943
 DB 2042 GGGCGGACGACGATCCAGCCACCCGCTCTCGTCGACACGATCGAAGAAGAT 2101
 QY 2944 CAAGCACGCTCGGCCATCCCGGAGCGCAAGAAGATCTCCAGGAGAGGAGGAGCG 3003
 DB 2102 GATTGAGACCTCTCCCGTTCCCGTCCGCAAAACAGC-----AGACGACGAGAGAA 2152
 QY 3004 CAAGGAGTTCCGCATCAGGAGTACTCGCCGCAAGCCTTCAAGTTCTTCATGGAGCAGCA 3063
 DB 2153 GAGAGAGTCACGACATCAAGAGCTACTCGCCATACGCCCTTTAAGTTCTTCATGGAGCAGCA 2212
 QY 3054 CATAGAGAACGTCATCAAGTCGTCGACGACGATCGCAAGATCAAGTGGAGAA 3123
 DB 2213 CGTGAGAGTGTATCAAAACCTACGACGACGAGTAAACCGAGGCTGACGTCGGAGCA 2272
 QY 3124 GGAGATGACAAAGTGGGACTGCCCGATCGACACCAATCGAGATGAGGAAATGCTGAA 3183
 DB 2273 AGAAATGCCAAAGCTGGACTCTGTGAAGCTGAGCAGGACGATCGGAAGATCTCTTA 2332
 QY 3184 CCAAAGGAGACGACATACATTCGATTCGATTCGACGCGCCCAAGATGACAGACATGTTGCT 3243
 DB 2333 CCAGAAAGAGTCTAATACAAAGGTTAAAGAGGCGCAAGATGACAGTCTATGTTGT 2392
 QY 3244 CAAACTGAAGCCCATTTGGGAGTGGTGCATTTGGCGAGTTAACGCTGCTGAGCAAAATCGA 3303
 DB 2393 CAAGATCAAAACCCCTGGGATCGGTCGCTTTGGAGAGTGTGCTTCTGTAAGTGGGA 2452
 QY 3304 TACCTCAACCATTTGATGGATGAAACCTCGGAAAGCGGACGCTTCTCAAGCCGAA 3363
 DB 2453 ---CACTACGCGCTGTACGCCATGAAGACCTAAGGAAAGAGATGCTCTGAACCCGAA 2509
 QY 3364 TCAGTGGCACACGTCGAGGAGGATATCTCGCGGAGGCGGACATCTCGCGGAGGCGGACAT 3423
 DB 2510 TCAGTGGCGGACGTCGAGGCGGAGGAGGACATCTCGCGGAGGCGGACATGAGTGGGT 2569
 QY 3424 GTGAAGTTGTACTACGCTTCCAGGACAAAGGATAATCTGTTGTTGTTGAGTACAT 3483
 DB 2570 GGTCAAACTCTACTACTCTTCCAAAGACAAAGACAGCTCTGTTGTTGAGTACAT 2629
 QY 3484 ACCAGTGGTGTATGTCGCTGTCATCAAACTGGGCTATTTTCGAGGAGGAGTGGC 3543
 DB 2630 CCCTGGTGGGACATGATGAGCCTGTGATCCGGATGGAGGTCTCCCTGAGCACCTGGC 2689
 QY 3544 CAGATTCTACATCGCGGAGGTCACCTCGCGCTGGAGGAGGTTCACAAAATGGCTTCAT 3603
 DB 2690 CCGGTTCTACATCGCAGAGTGTGCTTGGCCATGAGAGTGTCCCAAGATGGCTTCAT 2749
 QY 3604 TCACAGAGACATCAAGCCTGACAACTATCTCATGATGAGGACGACACATAAGCTCAC 3663
 DB 2750 CCACCGAGACATCAAGCCTGATTAACATTTGATAGATCTGATGGTTCACATTAACCTCAC 2809
 QY 3664 CGACTTTGGCTCTGCACGGGATTCGATGGAGGACGACCACTCGAAAGTACTACAGAGAA 3722
 DB 2810 AGATTTCGGCTCTGCACTGGGTTTCAGGTGAGTCAATTCACAAATATATACCAAGAGG 2869
 QY 3723 -----ACGCAATCACTCGCGCCAGGACTCGATGGAGGCTCGGAGGAGTATACCGAA 3774
 DB 2870 GAGCCATGTACAGACGACAGCATGAGCCACGACCTCTGGGATGATGTTGTTACTG 2929
 QY 3775 ---GACGACCGAAGCCACCTGCTGAGGAGGCGGAGGATCGCGGATCAACAAAGAT 3831
 DB 2930 TCGGTGTGGGACGAGGCTGAAGACCCCTAGAGCAGAGGCGGCGGAGCAGCAGAGGTG 2989
 QY 3832 CTTGGCCCACTCTGCTGTGGGACCCCGCACTACATAGTCTCCGAGGTCTCTGAGAGRAG 3891
 DB 2990 CTTGGCACATCTACTGTGGGGACTCAAACTACATCGCACCCGCGGAGGTGCTCTCCGCAA 3049

OS	Homo sapiens.
XX	
FH	Key
FT	Location/Qualifiers
CDS	2..1498
FT	*tag= a
FT	/product= "kinase"
FT	
PN	W0200008180-A2.
XX	
PD	17-FEB-2000.
XX	
PF	03-AUG-1999; 99WO-US17630.
XX	
PR	04-AUG-1998; 98US-0095270.
PR	11-SEP-1998; 98US-0099972.
PA	(IMMV) IMMUNEX CORP.
XX	
PI	Virca GD, Bird TA, Anderson DM, Marken JS;
XX	
DR	WPI; 2000-195584/17.
DR	P-PSDB; AAY69160.
XX	
PT	New human kinase polypeptides and polynucleotides used as molecular
PT	weight markers and as controls for peptide fragmentation -
XX	
PS	Claim 1; Page 8-9; 60pp; English.
XX	
CC	The present sequence encodes a polypeptide which has kinase
CC	activity. The kinase polynucleotides can be used to express the
CC	polypeptides, and as probes to identify nucleic acids encoding
CC	proteins having kinase activity. The kinase polypeptides and
CC	fragmented polypeptides are used as molecular weight and isoelectric
CC	focusing markers, and as controls for peptide fragmentation. They also
CC	have a number of therapeutic uses as kinases play a central role in
CC	cellular signal transduction. The polypeptides could also be used to
CC	identify binding partner proteins. The polypeptides can also be used as
CC	a reagent to identify any proteins that the polypeptide regulates, and
CC	proteins with which it might interact. The polypeptides may also be
CC	used for preparation of antibodies. The antibodies can be used in
CC	assays to detect the presence of the protein, and to purify the protein
CC	by immunoaffinity chromatography.
XX	
SQ	Sequence 1498 BP; 384 A; 417 C; 427 G; 270 T; 0 other;
Query Match. 9.4%; Score 536; DB 21; Length 1498;	
Best Local Similarity 65.5%; Pred. No. 2.4e-103;	
Matches 884; Conservative 3; Mismatches 433; Indels 30; Gaps	
QY	2884 GGCGGCAGCAAGGATCCACGGCACCACGCCCTCTCTCGTCCAGCACGTCGAAGAAGAT 2943
Dd	
Dd	160 GGCGCGCGACAAGAGCGCGAANAAGCGCCAAAGGGGGACAAAAGGGGAAGGATAAAGCA 219
QY	2944 CAAGCACGCGCTCCGCCCATCCCGGAGCGAAGAAGATCTCCAAGGAGGAGGAGGAGCG 3003
Dd	
Dd	220 GATTACAGACTCTCCGTTCGCCGTCCGCAAAAACAGC-----AGAGACGAAGAA 270
QY	3004 CAAGGAGTTCCGATCAGGCGAGTACTCGCGCGCAAGCCCTTCAAGTTCTTCATGGAGACGA 3063
Dd	
Dd	271 GAGAGAGTCAAGCATCAAGAGTACTTCGCGCATACGCCCTTTTAAGTTCTTCATGGAGACGA 330
QY	3064 CATAGAGACGGTGATCAAGTCTATCGCCAGCGCAGGTATCGCAAGAATCAGCTGGAGAA 3123
Dd	
Dd	331 CGTGGAGAATGTATCAAACCTTACAGCAAGGTTAACCGGAGGCTGCAGCTGGAGCA 390
QY	3124 GGAGATGCACAAAGTGGGACTCCGCCGATCAGACCCCAAAATTCAGATGAGGAAATGCTGAA 3183
Dd	
Dd	391 AGAAATGGCAAGCTGGACTCTGTGAAGCTGAGCAGGAGCAGATGCGGAAGATCCTCTA 450
QY	3184 CCAAAGGAGGACCACTACTCATTTGATTGAAGCGCGCCCAAGATGGCAAGAGCATGTTCTGT 3243
Dd	
Dd	451 CCAGAAGAGTCTTAATTACCAAGAGTTAAAGAGGCGCAAGATGGACAAGTCTATGTTGT 510

XX
KW Kinase activity; molecular weight marker; isoelectric focusing marker;
KW peptide fragmentation control; cellular signal transduction; ss.

Qy 3244 CAAACTGAAGCCCATTTGGAGTGGGTGATTTGGCGAGGTAAACGCTGTGTGAGCAAAAATCGA 3303
Db 511 CAAGATCAAAACCCCTGGGATCGGTGCTTTGGAGAAGTGTGCCTTCTTAAGTGA 570
Qy 3304 TACCTCAACACATTTGATGGGTGAARACCTTCGCGAAAGCGACCTTCTCAAGCGGAA 3363
Db 571 ---CACTCAGCCCTGTACGCCATGAGACCCCTAAGAAAAGGATGCTCTGAACCGGAA 627
Qy 3364 TCAGTGGCCACACGTGAAGCGGAGAGGATATCTCCGCGAAGCGCACATAACTGGGT 3423
Db 628 TCAGTGGCCACACGTGAAGCGGAGAGGATATCTCCGCGAAGCGCACATGAGTGGGT 687
Qy 3424 GGTGAAGTTGATACAGCTTCCAGGACAAGGATAATCTGTACTTTGTATGGACTACAT 3483
Db 688 GGTCAAACTCTACTACTCTCCAGACAAGACAGCCTGTACTTTGTATGGACTACAT 747
Qy 3484 ACCAGTGGTGTATCTGATGCTGCTCATCAAACTGGGCATTTTCGAGGAGGAACCTGGC 3543
Db 748 CCCTGGTGGGACATGATGAGCCTGCTGATCCGGATGAGGTCTTCCCTGAGCACCTGGC 807
Qy 3544 CAGATTCTACATCCCGAGGTACCTGCGCGTGACAGAGGTTTCACAAAATGGGCTTCAT 3603
Db 808 CCGGTTCTACATCCGAGAGTACTTTGGCCATTGAGAGTGTCCACAAGATGGGCTTCAT 867
Qy 3604 TCACAGAGACATCAAGCCTGACAACTACTCATCATGATAGGACGCGACACATAAGCTCAC 3663
Db 868 CCACCGAGACATCAAGCCTGATAAATTTGATAGATCTGGATGGTGTACATTAACCTCAC 927
Qy 3664 CGACTTTGGCTGTGACGGGATTCGATGAGCAGCAAACTCGAAGTACTACACAGGA--- 3720
Db 928 AGATTTCGGCCTGCACTGGGTTTCAGGTGGACTCAAAATTTCCAAATATTACCAAGAAGG 987
Qy 3721 ---GAACGGCAATCACTCGCCGAGGACTCGATGGAGCCCTGGGAGGA-----ATACTC 3771
Db 988 GAGCATGTACAGACAGACACATGAGGCCCGACCTCTGGGATGTGTCTAATG 1047
Qy 3772 CGAAGAGGACCGAAGCCACCGTGTGGAGAGGCGGAGTGGCGATCACCAAGAGT 3831
Db 1048 TCGGTGTGGGACAGGCTGAAGACCTAGACAGAGGCGCGGAAGCAGCAGAGAGTG 1107
Qy 3832 CCTGGCCACTCGTGTGGGACCGGACCTACATAGTCCCGAGTGTGGAGAGAG 3891
Db 1108 CTGGCACATTCATGTGGGAGCTCCAACTATATCGCACCGAGGTGCTCTCCCGAA 1167
Qy 3892 TGGTACACGAGCTGTGCGACTACTGAGCGTGGCGCTCATCTCTTAYGAGATGCTGT 3951
Db 1168 AGGTACACTCAACTCTGTGACTGTGGAGTGTGGAGTGATTTCTTCGAGATGCTGT 1227
Qy 3952 GGTTCAGCGCCCTTTCTGGCCACAGTCCGCTGGAACGCAACAAAGGTCACTCACTG 4011
Db 1228 GGGCAGCGCCCTTTTGGCACCTACTCCACAGAACCCAGCTGAAGGTGATCACTG 1287
Qy 4012 GGAGAAACCTGATATTCCCGCGAGCGGAGTATCCCGGAGGCTACGGACTTGAT 4071
Db 1288 GGAGAACAGCTCCACATTCAGCCAGGTGAAGCTGAGCCCTGAGGCGGAGGACCTCAT 1347
Qy 4072 AAGGAGCTGTGCTGCTGCGTGAACAGCGCTGGGCAAGAGCGTGGACAGG---TCAA 4128
Db 1348 CACCAAGCTGTGCTCTCCGAGACACCGCCTGGGCGGGAATGGGCGGATGACCTGAA 1407
Qy 4129 GAGCCAGACTTTCTTAAGGCGATCGACTTTGC---GGACATGGGGAAGCAAGCGCC 4185
Db 1408 GGGCAGCCCTTTCTCAGCGCCTTGACTTCTTCAGTGTGATCTCCAGTGCATCCGGAAGCATCCAGCGCC 1467
Qy 4186 CTACATCCGGAATCAAGCACCCACRACGA 4215
Db 1468 CTAGTTCACCATCAGCCACCCCATGGA 1497

RESULT 13

ABK43457

ID ABK43457 standard; cDNA; 1912 BP.

XX

AC ABK43457;
XX 05-JUN-2002 (first entry)
DT
XX
DE DNA encoding novel central nervous system protein #37.
XX
XX
KW Central nervous system; CNS; autoimmune disease; rheumatoid arthritis;
KW hyperproliferative disorder; neoplasm; cardiovascular disorder;
KW cardiac arrest; cerebrovascular disorder; ischemia; angiogenesis;
KW nervous system disorder; Alzheimer's disease; AIDS; ocular disorder;
KW acquired immunodeficiency virus; dysphagia; gastrointestinal disorder;
KW adenocarcinoma; reproductive system disorder; testicular feminisation;
KW endocrine disorder; diabetes; cancer; leukaemia; neovascularisation;
KW respiratory disorder; renal disorder; kidney failure; blood disorder;
KW myocardial infarction; wound healing; cell proliferation; skin aging;
KW food additive; food preservative; gene therapy; gene; ss.
XX Homo sapiens.
OS
XX
XX WO200155318-A2.
PN
XX
XX 02-AUG-2001.
XX
XX 17-JAN-2001; 2001WO-US01332.
XX
XX 31-JAN-2000; 2000US-0179065.
PR 04-FEB-2000; 2000US-0180628.
PR 24-FEB-2000; 2000US-0184664.
PR 02-MAR-2000; 2000US-0186350.
PR 16-MAR-2000; 2000US-0189874.
PR 17-MAR-2000; 2000US-0190076.
PR 18-APR-2000; 2000US-0198123.
PR 19-MAY-2000; 2000US-0205515.
PR 07-JUN-2000; 2000US-0209467.
PR 28-JUN-2000; 2000US-0214886.
PR 30-JUN-2000; 2000US-0215135.
PR 07-JUL-2000; 2000US-0216847.
PR 07-JUL-2000; 2000US-0216880.
PR 11-JUL-2000; 2000US-0217487.
PR 11-JUL-2000; 2000US-0217496.
PR 14-JUL-2000; 2000US-0218290.
PR 26-JUL-2000; 2000US-0220963.
PR 26-JUL-2000; 2000US-0220964.
PR 14-AUG-2000; 2000US-0224518.
PR 14-AUG-2000; 2000US-0224519.
PR 14-AUG-2000; 2000US-0225213.
PR 14-AUG-2000; 2000US-0225214.
PR 14-AUG-2000; 2000US-0225266.
PR 14-AUG-2000; 2000US-0225267.
PR 14-AUG-2000; 2000US-0225268.
PR 14-AUG-2000; 2000US-0225270.
PR 14-AUG-2000; 2000US-0225447.
PR 14-AUG-2000; 2000US-0225757.
PR 14-AUG-2000; 2000US-0225758.
PR 14-AUG-2000; 2000US-0225759.
PR 18-AUG-2000; 2000US-0226279.
PR 22-AUG-2000; 2000US-0226691.
PR 22-AUG-2000; 2000US-0226698.
PR 22-AUG-2000; 2000US-0227182.
PR 23-AUG-2000; 2000US-0227009.
PR 30-AUG-2000; 2000US-0228924.
PR 01-SEP-2000; 2000US-0229287.
PR 01-SEP-2000; 2000US-0229343.
PR 01-SEP-2000; 2000US-0229344.
PR 01-SEP-2000; 2000US-0229345.
PR 05-SEP-2000; 2000US-0229509.
PR 05-SEP-2000; 2000US-0229513.
PR 06-SEP-2000; 2000US-0230437.
PR 06-SEP-2000; 2000US-0230438.
PR 08-SEP-2000; 2000US-0231242.
PR 08-SEP-2000; 2000US-0231243.
PR 08-SEP-2000; 2000US-0231244.
PR 08-SEP-2000; 2000US-0231413.

PR	08-SEP-2000;	2000US-02314111;
PR	08-SEP-2000;	2000US-02320801;
PR	08-SEP-2000;	2000US-02320801;
PR	12-SEP-2000;	2000US-02319681;
PR	14-SEP-2000;	2000US-02323971;
PR	14-SEP-2000;	2000US-02333981;
PR	14-SEP-2000;	2000US-02333991;
PR	14-SEP-2000;	2000US-02334001;
PR	14-SEP-2000;	2000US-02340111;
PR	14-SEP-2000;	2000US-02330631;
PR	14-SEP-2000;	2000US-02330641;
PR	14-SEP-2000;	2000US-02330651;
PR	21-SEP-2000;	2000US-02344223;
PR	21-SEP-2000;	2000US-02344274;
PR	25-SEP-2000;	2000US-02349971;
PR	25-SEP-2000;	2000US-02349981;
PR	25-SEP-2000;	2000US-02349981;
PR	26-SEP-2000;	2000US-02345484;
PR	27-SEP-2000;	2000US-02338634;
PR	27-SEP-2000;	2000US-02338633;
PR	27-SEP-2000;	2000US-02338632;
PR	29-SEP-2000;	2000US-02363671;
PR	29-SEP-2000;	2000US-02363681;
PR	29-SEP-2000;	2000US-02363691;
PR	29-SEP-2000;	2000US-02363701;
PR	02-OCT-2000;	2000US-02368602;
PR	02-OCT-2000;	2000US-02370371;
PR	02-OCT-2000;	2000US-02370381;
PR	02-OCT-2000;	2000US-02370381;
PR	02-OCT-2000;	2000US-02370391;
PR	02-OCT-2000;	2000US-02370401;
PR	13-OCT-2000;	2000US-02399935;
PR	13-OCT-2000;	2000US-02399937;
PR	20-OCT-2000;	2000US-02409660;
PR	20-OCT-2000;	2000US-02411221;
PR	20-OCT-2000;	2000US-02417851;
PR	20-OCT-2000;	2000US-02417861;
PR	20-OCT-2000;	2000US-02418081;
PR	20-OCT-2000;	2000US-02418091;
PR	20-OCT-2000;	2000US-02418261;
PR	01-NOV-2000;	2000US-02446171;
PR	08-NOV-2000;	2000US-02464741;
PR	08-NOV-2000;	2000US-02464751;
PR	08-NOV-2000;	2000US-02464761;
PR	08-NOV-2000;	2000US-02464771;
PR	08-NOV-2000;	2000US-02464781;
PR	08-NOV-2000;	2000US-02465231;
PR	08-NOV-2000;	2000US-02465241;
PR	08-NOV-2000;	2000US-02465251;
PR	08-NOV-2000;	2000US-02465261;
PR	08-NOV-2000;	2000US-02465271;
PR	08-NOV-2000;	2000US-02465281;
PR	08-NOV-2000;	2000US-02465321;
PR	08-NOV-2000;	2000US-02466091;
PR	08-NOV-2000;	2000US-02466101;
PR	08-NOV-2000;	2000US-02466111;
PR	08-NOV-2000;	2000US-02466131;
PR	17-NOV-2000;	2000US-02492107;
PR	17-NOV-2000;	2000US-02492081;
PR	17-NOV-2000;	2000US-02492091;
PR	17-NOV-2000;	2000US-02492101;
PR	17-NOV-2000;	2000US-02492111;
PR	17-NOV-2000;	2000US-02492121;
PR	17-NOV-2000;	2000US-02492131;
PR	17-NOV-2000;	2000US-02492141;
PR	17-NOV-2000;	2000US-02492151;
PR	17-NOV-2000;	2000US-02492161;
PR	17-NOV-2000;	2000US-02492171;
PR	17-NOV-2000;	2000US-02492181;
PR	17-NOV-2000;	2000US-02492191;
PR	17-NOV-2000;	2000US-02492201;
PR	17-NOV-2000;	2000US-02492211;
PR	17-NOV-2000;	2000US-02492221;
PR	17-NOV-2000;	2000US-02492231;
PR	17-NOV-2000;	2000US-02492241;
PR	17-NOV-2000;	2000US-02492251;
PR	17-NOV-2000;	2000US-02492261;
PR	17-NOV-2000;	2000US-02492271;

PR	17-NOV-2000;	2000US-02492999;
PR	01-NOV-2000;	2000US-02493000;
PR	17-DEC-2000;	2000US-0250160;
PR	01-DEC-2000;	2000US-0250391;
PR	05-DEC-2000;	2000US-0251030;
PR	05-DEC-2000;	2000US-0251988;
PR	05-DEC-2000;	2000US-0256719;
PR	06-DEC-2000;	2000US-0251479;
PR	08-DEC-2000;	2000US-0251856;
PR	08-DEC-2000;	2000US-0251868;
PR	08-DEC-2000;	2000US-0251869;
PR	08-DEC-2000;	2000US-0251989;
PR	08-DEC-2000;	2000US-0251990;
PR	11-DEC-2000;	2000US-0254097;
PR	05-JAN-2001;	2001US-0259678;

(HUMA-) HUMAN GENOME SCI INC.

Rosen CA, Barash SC, Ruben SM;

WPI; 2001-581633/65.

P-PSDB: AAU87127.

New isolated nucleic acid encoding a protein for diagnosing, preventing, treating or ameliorating medical conditions and used as food additives or preservatives -

Claim 1; SEQ ID No 47; 837pp; English.

The invention describes an isolated nucleic acid molecule (I) encoding a novel central nervous system protein. (I) and polypeptides (III) encoded by (I), are used to treat a medical condition and in diagnosis of a pathological condition. Disorders which are diagnosed or treated include autoimmune diseases e.g. rheumatoid arthritis, hyperproliferative disorders e.g. neoplasms of the breast or liver, cardiovascular disorders e.g. cardiac arrest, cerebrovascular disorders e.g. cerebral ischaemia, angiogenesis, nervous system disorders e.g. Alzheimer's disease and amyotrophic lateral sclerosis, infections caused by bacteria, viruses e.g. Acquired immunodeficiency virus (AIDS) and fungi, ocular disorders e.g. corneal infection, gastrointestinal disorders e.g. dysphagia, adenocarcinomas and irritable bowel syndrome, reproductive system disorders e.g. testicular feminisation, endocrine disorders e.g. diabetes and pituitary dwarfism, cancers and disorders at the cellular level e.g. leukaemia, disorders involving neovascularisation e.g. malignancies, respiratory disorders e.g. nonallergic rhinitis, renal disorders e.g. acute kidney failure and blood related disorders e.g. myocardial infarction. The polypeptides can also be used to aid wound healing and epithelial cell proliferation, to prevent skin aging due to sunburn, to maintain organs before transplantation, for supporting cell culture of primary tissues, to regenerate tissues and in chemotaxis. The polypeptides can also be used as a food additive or preservative to increase or decrease storage capabilities, fat content, lipid, protein,

Query Match 9.4%; Score 535.2; DB 23; Length 1912;
Best Local Similarity 65.5%; Pred. NO. 4e-103;

Best local similarity	Freq. NO.	4E-103
Matches 882; Conservative	4; Mismatches 427;	Indels 33; Gaps 6;

QY 3041 TTCAAGTTCTTCATGGAGCAGCACATAGAGAACGTGATCAAGTCGTATCGCCAGCGCAGC 3100

Db
1 TTTAAGTTCCTCATGGAGCAGCAGTGGAGAATGTCAATAAACTACCAGCAGAAGGTT 60

3101 TATCGCAAGAATCAGGCTGGAGAGAGGAGATGCACAAAGTGGGACTGCCCCGATCAGACCCCA 3160

61 AACCGAGGCTGCAGCTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAG 120

3161 ATCGAGATGAGGAAATGCTGAACCAAAAGGAGAGCAACTACATTCTGATTGAAGCGCGCC 3220

Db 121 GAGCAGATCGGGAAGATCCTCTACCCAGAAACAGCTCTAATTACAACAGGTTAAGAGGGCC 180

0v 3221 AAGATGGACAAGAGCATGTTCTCGTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAG 3280

181 A G A T T C G A C A A G T C T A T G C T T T C T C A C A C A T C A A A C C C T C C G C A T C C C T C C T T T C C G C A A 240

QY 3281 GTACGCTGTGACCAAAATCGATACCTCGAACCAATTTGTATGGGATGAAACCCCTCGG 3340
DB 241 GTGTGCTTGTGTTGAAGTGA---CACTCACGCCCTGACGCCATCAAGACCCCTAAG 297
QY 3341 AAAGCGGACGTTCTCAAGCGGAATCAGGTGCGCACACGTCGAAGCGCGAGGATATCCTC 3400
DB 298 AAAAAGATGTCCTGAACCGGAATCAGGTGCGCCACCTCAAGCGCGAGGAGCATCTG 357
QY 3401 CGGAAGCCGACATAACTGCGTGGTGAAGTGTACTACAGCTTCCAGGCAAGATAAT 3460
DB 358 GCCGAGCGACACATGAGTGGTGGTCAAACTCTACTCTCCATCCAAAGACAGCAGC 417
QY 3461 CTGTACTTTGTGATGAGTACATACAGTGGTGTATCTGATGCTGCTGCTCATCAAACTG 3520
DB 418 CTGTACTTTGTGATGAGTACATACCTGCTGGTGGGACATGATGAGCTGCTGATCCGATG 477
QY 3521 GGCAATTTTCGAGGAGGAGTGGCGAGATTCTACATCCCGGAGTCACTGCGCGGTGAC 3580
DB 478 GAGTCTTCCCTGAGCACCTGGCCCGGCTTCTACATCGCAGAGCTGACTTTGGCCATTTGAG 537
QY 3581 AGCGTTCCAAAATGGCTTCATTCACAGAGACATCAAGCTGACAACTACTCATCGAT 3640
DB 538 AGTGTCCACAAGATGGCTTCATCCACCGAGACATCAAGCTGATACATTTTCATAGAT 597
QY 3641 AGGACGGACACATAAAGCTACCGACTTTGGCTGTGACGAGGATTCGATGGACGCAC 3700
DB 598 CTGATGCTCACATTAACCTCACAGATTTCGGCTCTGCACATGGTGGTTCAGTGGACTCAC 657
QY 3701 AACTCGAAGTACTACGAGGAGACGGCAATCACTCGC-----GCCAGGACTCGATGGAG 3754
DB 658 AATTCCAATATTACCAGAAGGAGGCATGTACAGACGAGCAGCATGGAGCCCGACGCAC 717
QY 3755 CCCTGGGAGGAAT-----ACTCCGAGAACCGACCGACCGCTGCTGGAGAGCGCA 3808
DB 718 CTCTGGATGATGTCTACTGTCTGCTGGTGGGACAGGCTGAGACCCCTAGACGAGG 777
QY 3809 CGGATCGCGATCACCAAGAGTCTTGGCCACTCGCTGTGGGACCCCGCAACTACATA 3868
DB 778 GCGGGAAGCAGCAGCAGAGGCTGCTGGCACATTCATCTGTGGGAGTCCAAACTACATC 837
QY 3869 GCTCCCGAGTGTGGAGAGTGGTACACGAGCTGTGGCACTGTGGCACTGAGCGGTGGC 3928
DB 838 GCACCCGAGTGTCTTCCGCAAGGATACATCACTGTGACTGTGGTGGAGTGTGA 897
QY 3929 GTATCTCTTAYGAGATGCTGGTGGTGCAGCGCCCTTTCTGGCAACAGCTCCGCTGGAA 3988
DB 898 GTGATCTCTTCGAGATGCTGGTGGGACCGCCCTTTTGGCACTTCTCCACAGAA 957
QY 3989 ACCGACAAAAGTCACTAACTGGGAGAAACSTGCGATATTCGCGGAGCGGAGTTA 4048
DB 958 ACCGAGCTGAAGTGTATCAACTGGGAGAACAGCTCCACATTCAGCCAGGTGAAGCTG 1017
QY 4049 TCCGCGGAGGCTACGAGTGTAAAGGAGGCTGTGCGTCGCTGACAGCGGCTGGC 4108
DB 1018 AGCCCTGAGGCGGAGGACCTCATCACCAGCTGTGCTGCCGACAGCACCCGCTGGG 1077
QY 4109 AAGAGCGTGGACGAGG---TCAAGAGCCAGGACTTCTTCAAGGCGCATGCACTTTCG---G 4162
DB 1078 CGGAATGGGCGGATGACCTGAGAGGCCACCCCTTCTCAGCGCCATGACTTCTCCAGT 1137
QY 4163 GACATCGGAGGAGAAAGCGCCCTCATATCCGGAATCAAGCAACCCACGAGCACATCC 4222
DB 1138 GACATCGGAGGAGGAGCGCCCTACGTTCCACCACCATGACGACCCCGCATGAGCACTCG 1197
QY 4223 AACTTTGATCCCTGATCGGAGAGGCTGCTGCAATGACTCCACCATGACGAGCGGC 4282
DB 1198 AATTTGACCCCTGATGATGAAGAAGCCCTTGGAGACATGCCAGGAGGATGACCAAG 1257
QY 4283 GATGATGTC-----GACCAGATGACCGGACATCTTCACGCGCTTTTTCGAAATTT 4330
DB 1258 GCCTGGGACACACTCACTCGCCCAATAAGACATCCTTGAGCAACGCAATTTACGAAATTC 1317
QY 4331 ACCTTCGCTGCTTCGACGACAA 4356

DB 1318 ACCTTCGAAGGTTCTTTTGATGACAA 1343

RESULT 14

AA42119
ID AAT42119 standard; cDNA; 3213 BP.

XX AAT42119;

XX 31-JAN-1997 (first entry)

XX M-lats gene encoding large tumour suppressor.

XX Mouse; m-lats gene; large tumour suppressor; fetal brain;
protein-serine/threonine-kinase; cell proliferation; antisense;
dominant-negative; cancer; degenerative disorder; trauma;
growth deficiency; therapy; antitumour; vulnary; diagnostic;
transgenic plant; transgenic animal; growth; senescence; ds.

XX Mus musculus.

XX Key Location/Qualifiers
CDS I..2889

FT /*tag= a
FT /product= m-lats protein

XX WO9630402-A1.

XX 03-OCT-1996.

XX 26-MAR-1996; 96WO-US04101.

XX 27-MAR-1995; 95US-0411111.

XX (UYUA) UNIV YALE.

XX Tao W, Wang W, Xu T, Yu W, Zhang S;

XX WPI; 1996-455275/45.

XX P-PSDB; AAW05179.

XX New isolated large tumour suppressor gene - used to develop prods.
for inhibiting cell proliferation or for enhancing proliferation

XX Disclosure; Page 126-130; 215pp; English.

XX This sequence encodes a mouse large tumour suppressor m-lats protein,
and has been isolated from a newborn mouse brain phage lambda-ZAP
cDNA library using a 2.2-kb DNA probe from the Drosophila lats gene
(AAT42117). A homologous mouse sequence has also been isolated
(m-lats2, AAT42120). The gene encodes a putative protein-
serine/threonine-kinase, and inhibits cell proliferation and plays a
crucial role throughout development. Activators or inhibitors of
lats function (e.g. an antisense oligonucleotide or dominant-negative
lats fragment) may be used in therapy of cancer or other
proliferative disorders, degenerative disorders, trauma, growth
deficiency, etc., and fragments of the gene may be used as diagnostic
probes. A lats-inhibitor sequence may be expressed in a transgenic
plant or farm animal to confer increased growth and inhibit
senescence.

XX Sequence 3213 BP; 946 A; 762 C; 697 G; 808 T; 0 other;

Query Match 9.1%; Score 522.8; DB 17; Length 3213;
Best Local Similarity 65.3%; Pred. No. 2.1e-100;
Matches 850; Conservative 4; Mismatches 421; Indels 27; Gaps 5;

QY 2984 AAGGAGAAGGAGGAGGAGCGCAAGGAGTTCGCGATCAGGCGAGTACTCGCGCAAGCCTTC 3043

DB 1354 AACAAGAAGATCAAGACGACGAGAGTCTCGGATTCAGAGTTACTCTCCACAGGCTTT 1413

QY 3044 AGTTCTTCATGAGGAGGAGCAGACATAGAGACGTCATCAAGTCGTATCGCCACGACGTAT 3103

Db 1414 AAGTTCTTATGGAGCAGCAGCAGTGTCTGAGTCTCTCATCAGCAGGCTGTGCAT 1473
Qy 3104 CCACAAGTACAGTGGAGAGGAGTGCACAAAGTGGGAGTCCCGGATCAGACCCCAAAATC 3163
Db 1474 CGAAGAGCAGCTAGAAATGAATGATGGGGTGGATTATCTCAAGATGCCAGGAT 1533
Qy 3164 GAGATGAGGAAATGCTGAACCAAGAGAGCAACTACATTCGATTGAAGCGCCCAAG 3223
Db 1534 CAAATGAGAAAGATGCTTTGCCAGAAAGAGTCTAACTATATTCGTTTAAAGGGCTTAA 1593
Qy 3224 ATGGACAGAGCATGTTGCTCAAACTGAAGCCCATTTGGAGTGGGTCATTTGGCGAGGTA 3283
Db 1594 ATGGCAAGTCTATGTTGTTAAAGATAAAGACATAGGAATAGAGCGGTTTGGTGAAGTC 1653
Qy 3284 AGCTGGTGAGCAAACTCGATCCCTCGAACCAATTTGTATGCGATGAAACCCCTCGGAA 3343
Db 1654 TGTCTAGCAAGAAAGTCTGATA--CTAAGCTTTGTATGCAACAAAGACTCTTGAAG 1710
Qy 3344 CGGAGCTTCTCAAGCGGAATCAGGTGGCAGACAGTGAAGCGGAGAGGATATCCTCGG 3403
Db 1711 AAAGAGTCTCTCGGAAATCAGGTGGCTCATGTGAAGCGGAGAGGATATCTAGCA 1770
Qy 3404 GAAGCCGACAAATACCTGGGTGGTGAAGTGTACTACAGCTTCCAGCAAGGATTAATCTG 3463
Db 1771 GAAGCCGACAAATGAGTGGGTGGTGGCTGTACTCTTCTCCAGCAAGGACAACTTG 1830
Qy 3464 TACTTTGTGATGGACTACATACACAGTGTGATCTGCTGCTCATCAAACTGGG 3523
Db 1831 TACTTTGTGATGGACTACATCTCTGGGGGATATGATGAGCTATTAATTAAGATGGG 1890
Qy 3524 ATTTTCGAGGAGAACTGGCCAGATTTACATCGCGGAGGTCACTGGCGCCGTGCACAGC 3583
Db 1891 ATCTTTCTGAAATCTGGCAGATTTACATAGCAGAACTTACCTGTGCACTTGAAGT 1950
Qy 3584 GTTCAAAATGGGCTTCATTCACAGACATCAAGCCGTGACAACTACTATCATAGG 3643
Db 1951 GTTCAAAATGGGTTTATTTCATAGATATTAACCTGTATGAATTTGATGACCGT 2010
Qy 3644 GACGACACATAAAGCTCACCGACTTTCGCTGTGCAGGGATTCGGATGGAGCGACAAAC 3703
Db 2011 GATGCCATATTAATTTACTGACTGTTGGCTTGCATCGCTTCCAGATGGACATGAC 2070
Qy 3704 TCGAGTACTACAGGAGAGGCAATCACTCGCCAGGACTCGATGAGCGCCCTGGAG 3763
Db 2071 TCCAGTACTACC---AGAGTGGGATCAGCCAGCAAGATAGCATGATTTCAATAC 2127
Qy 3764 GAATPACTCGA-----GAAGGACCGAGCCACCGCTGTGGAGGGCGA 3808
Db 2128 GAATGGGAGATCTTCCAAATGTGCGGTGGGACAGACTGAAGCCACTGGAGCGGAGA 2187
Qy 3809 CGGATGCGGATCACCAAGAGTCTGCGCCACTCGCTGGTGGGACCCCGCAACTACATA 3868
Db 2188 GCTGCTCCGACACACCGGATGTAGCCCAATCTCTGGTGGGACTCCCAATTAAT 2247
Qy 3869 GCTCCGAGGTGCTGGAGAGTGGGTACACGAGCTGTGCGACTACTGGAGCGTGGGC 3928
Db 2248 GCACCTGAAGTCTACTCGCAAGAGATATACACAGCTGTGTGACTGGTGGAGTGTGT 2307
Qy 3929 GTCATCTTAYAGATGCTGTGGTGGTGGGAGCGGCTTTCGGCCCAAGTGGCTGGAA 3988
Db 2308 GTTATCTTGTGAATTTGTGGGACAACTCTCTTCTTGGGACAAACCCCATTAGAA 2367
Qy 3989 ACGCAAAAAGGTCATCACTGGGAGAAACSCCTGCATATTCGCGCCGAGCCAGTTA 4048
Db 2368 ACACAAATGAAGTTATCTATGGCAAACTCTCTACACATCCCTTCTCAAGCTAAGCTG 2427
Qy 4049 TCCCGAGGTACTGAGCTTGTATAGGAGGTCTGTGGCTGGCTGCAAGCGGTGGGC 4108
Db 2428 AGTCTGAAGCTCTGACCTTATTAACATGTGTCCAGGACCAAGACCGCTCGGC 2487
Qy 4109 AAGA---CGGTGGAGGAGTCAAGAGCCAGCACTTCTTCAAGGGCATCGA---CTTTCG 4162
|||||
|||||

Db 2488 AAGAACGGTGTGTGAGATAAAGGCTCATCCATTTTAAAGACCATGATTCTCTAGT 2547
Qy 4163 GACATGCGGAGCAAGAGCGCCCTACATACCGGAAATCAAGCACCCACGACATCC 4222
Db 2548 GATCTGAGACAGATGCTGCTTCATACATCCCTAAATCACCATCCACAGATACATCC 2607
Qy 4223 AACTTTGATCCGTGGATCCGAGAGAGTCTGCGCTCGAATGAC 4264
Db 2608 AATTTGCGACCTGTTGATCTCTGATAAATTTGTGGAGGATGGC 2649

RESULT 15
AAZ51506

ID AAZ51506 standard; DNA; 3213 BP.

AC AAZ51506;

DT 21-JUN-2000 (first entry)

XX Mouse Lats (large tumour suppressor) DNA.

XX Mouse; Lats; large tumour suppressor; cytostatic; vulnery;
cell overproliferation inhibitor; cdc2; cell cycle-dependent kinase;
treatment; prevention; screening; cancer; skin; ovarian tumour;
soft tissue sarcoma; pituitary disorder; gene therapy; hyperplasia;
LH; luteinizing hormone hypogonadotropic hypogonadism; metaplasia;
dysplasia; degenerative disorder; growth deficiency; physical trauma;
hypoproliferative disorder; lesion; wound; lats knock-out mouse; ds.

OS Mus musculus.

XX Key Location/Qualifiers

FH 1..2889

FT **tag= a

FT /product= "Lats protein"

FT /partial

XX WO200010602-A1.

XX 02-MAR-2000.

XX 18-AUG-1999; 99WO-US19068.

XX 18-AUG-1998; 98US-0096996.

XX 18-AUG-1998; 98US-0096997.

XX (UIYA) UNIV YALB.

XX Xu T, Tao W, St John MAR, Fei X, Fukumoto RK, Zhang S;

XX Trenchalk GS, Stewart RA;

XX WPI: 2000-246496/21.

XX P-PSDB: AAY70391.

XX Use of lats proteins, complexes of lats and cdc2 for treating cancer
that is refractory to treatment by standard chemotherapy and radiation
therapy, and disorders associated with aberrant levels of cdc2 activity

PS Claim 44; Fig 13; 134pp; English.

XX The present sequence is a DNA encoding mouse Lats (large tumour
suppressor) protein which is a cell overproliferation inhibitor and a
negative regulator of cell cycle-dependent kinase cdc2/cyclin A.
XX The present sequence is useful for treating cancer that is refractory
to standard chemotherapy or radiation therapy such as hyperplasia,
metaplasia, or dysplasia, and disorders associated with aberrant
levels of cdc2 activity. Conditions treated by promoting cdc2 function
include degenerative disorders, growth deficiencies, hypoproliferative
disorders, physical trauma, lesions, and wounds. An animal model
preferably a mouse, in which a lats gene has been disrupted by homologous
recombination, e.g. a lats knock-out mouse, is used for screening
compounds that can be used to treat or prevent cancer, particularly

CC skin cancer, soft tissue sarcomas and ovarian tumours, and disorders
CC associated with pituitary dysfunction e.g. luteinizing hormone (LH)
CC hypogonadotropic hypogonadism. The lats DNA is also used in gene therapy.
xx

Sequence 3213 BP; 946 A; 764 C; 696 G; 807 T; 0 other;

Query Match.	9.18;	Score	522.8;	DB	21:	Length	3213;
Best Local Similarity	65.38;	Pred.	No. 2 le=100;				
Matches	850;	Conservative	4;	Mismatches	421;	Indels	27; Gaps
QY	2984	AAGGACAGGAGGAGCGCAAGGATCCGGATCAGGCAGTACTCGCGCGAAGCCTTC	3043				
Db	1354	ACACAGAAGATTGAAGAACGAAGAGAGTCTCGGATTCAGAGTTACTCCGCCACAGCCCTT	1413				
QY	3044	AAGTTCTTCATGGAGCAGCATAGAAAGCTGATCAAGTCGTATCGTATCGCACGCGCAGTAT	3103				
Db	1414	AAGTTCTTCATGGAGCAGCATAGAAAGCTGCTGAAGTCTATCAGCAGCGCTCGCAT	1473				
QY	3104	CCCAAGATCAGCTGGAGAGGAGTGCACAAAGTGGGACTGCCCGATCAGACCACCAATC	3163				
Db	1474	CGNAAGAGCAGCTAGAAAATGAAATGATCGCGGTTGGATTTATCTCAAGATGCCAGAT	1533				
QY	3164	GAGATGAGAAATGCTGAACCAAAGGAGAGCAACTACATTCGATTAAGGGCGCCAAG	3223				
Db	1534	CAATGACAAGATGCTTGCCGAAGAGCTAACATAATTTCGTCCTTAAAGSGCTAAA	1593				
QY	3224	ATGGACAAGACATGTTTCGTCACAACTCAAGCCCATTGGAGTCGGTGCAATTTGGCGCAGTA	3283				
Db	1594	ATGGACAAGTCTATGTTTGTAAAGATAAAGACATTAGGAATAGAGCGTTTGGTGAAGTC	1653				
QY	3284	ACGCTGGTGACAAAATCGATCCTCGAACCATTTGTATGGATGAAACCCCTCGGAAA	3343				
Db	1654	TGCTAGCAGAAGTCGATA---CTAAAGCTTTGTATGCAACAAAGACTCTTCGAAAG	1710				
QY	3344	CGGACGTTCTCAAGCGGAATCAGGTGGCACAAGTGAAGGGCGAGAGGATATCCTCGCG	3403				
Db	1711	AAAGACGTTCTGCTCCGAATCAGTGGCTCATGTGAAGCGGAGAGGATATCCTAGCA	1770				
QY	3404	GAACCCACAATAACTGGGTGGTGAAGTTGTACTACGCTTCCAGGACAAGGATTAATCTG	3463				
Ddb	1771	GAACCCACAATAAGTGGGTGGTGGCCGCTGTACTACTCTTCCAGGACAAGCAACTG	1830				
QY	3464	TACTTTGTGATGACTACATACAGTGGTGTATCTGATTCGGTGCATCAAACTGGCG	3523				
Db	1831	TACTTTGTGATGACTACATTCCTGGGGGGATATGATGACCTATTAAATAGATGGCG	1890				
QY	3524	ATTTTCAGGAGAACTGCCAGATTTACATCGCCGAGGTACCTGGCCGTTGGACAGC	3583				
Db	1891	ATCTTTCCTGAAATCTGGCAGATTTACATAGCAGAACTTACCTGTGCAGTTGAAAGT	1950				
QY	3584	GTTCAAAAATGGCTTCATTACAGAGACATCAAGCTGCACATACATCTATCGATAGG	3643				
Db	1951	GTTCAAAAATGGTTTTATTTCATAGAGATTTAACTGTATACTTTGATTGACCTT	2010				
QY	3644	GACGCACATTAAGCTCACCGCTTTGGCCTGTGCACGGGATTCGATGGAGCCGCAAC	3703				
Db	2011	GATGGCCATATTAAATTGACTGACTTTGGCTTTGTGCATCGCTTCAGATGGACATGAC	2070				
QY	3704	TCCAAGTACTACAGAGAACGGCAATCAGTCGCGCCAGGACTCGATGGAGCCCTGGGAG	3763				
Db	2071	TCCAAGTACTAC--AAGTGGGGATCACCCAGGCAAGATAGCATGATTTCAGTAAAC	2127				
QY	3764	GAATACTCCGA-----GAACGGACCGAAGCCCCACCTGCTGGAGAGCGCA	3808				
b	2128	GAATGGGAGATCTCTCCAAATTTGTCGGTGTGGGACAGACTGAAGCCACTGGAGCGGACA	2187				
QY	3809	CGGATGGCGGATACCAAGAGTCTTGCCCACTCGCTGTGGGGACACCGCAACTACATA	3868				
b	2188	GCTGCTGCCAGCACCGCGATGTCTAGCCCATCTCTGTTGGTGGGACTCCCAATTATATT	2247				
QY	3869	GTCGCCGAGGTCTGGAGAGRAGTGGGTACACGAGCTGTGCGCACTACTGCGAGCGTGGCG	3928				
b	2248	GCACCTGAAGTGCTACTCGCAACAGGATATACAGCTGTGTGACTGGTGGAGTGTGGT	2307				

Qy	3929	GTATCCTCTYTAGAGATGCTGGTGGCTCAGCGCGCCTTTCTGGCCAAACAGTCCGCTGGAA	3988
Db	2308	GTATATCTTGTGAATATGTTGGTGGGACAACTCCTTTCTGGCACAACCCCATAGAA	2367
Qy	3989	ACGCAACAAGAGTCACTAACTGGGAGAAAACTGCGATATTCCGCGCAGCGCGAGTTA	4048
Db	2368	ACACAATGAAGGTTCATCTGTGGCAAACTTCTCTACACATCCCTCCCTCAAGCTAAGCTG	2427
Qy	4049	TCCGCGGAGGCTACGGACATGTGATAGGAGGCTCTGTGGCTGGCTGACAGCGGCTGGCG	4108
Db	2428	AGTCCTTGAAGCCTCTCGACCTCATTTATCAACTGTGTCTGAGGACGAGAACGCGCTCGGC	2487
Qy	4109	AAGA---CGCTGACGAGGTCAAGAGCCACGACTTCTTCAAGGCGATCGA---CTTTGCG	4162
Db	2488	AAGAACGGTGTGATGAGATAAAGGCTCATCCATTTTTTAAGACCATCTGATTTCTCTAGT	2547
Qy	4163	GACATCGGAGACAGAAAGCGCCCTACATACCGGAAATCAAGCACCRCAGGACCATCC	4222
Db	2548	GATCTCGAGACAGCAGTCTGTTTCATACATCCCTAAATCAGCGATCCAAACAGATACATCC	2607
Qy	4223	AACTTTGATCCGCTGGATCCCGAGAGGCTGGCTCGAATGAC	4264
Db	2608	AATTCGACCCCTGTTGATCTCTGATAAAATGTGGAGCGATGCG	2649

Search completed: January 16, 2003, 10:36:52
Job time : 895.592 secs

GenCore version 5.1.3
Copyright (c) 1993 - 2003 Compugen Ltd.

OM nucleic - 'nucleic search, using sw model

Run On: January 16, 2003, 09:56:02 ; Search time 116.379 Seconds
(without alignments)
15073.116 Million cell updates/sec

Title: US-09-763-334-7
Perfect score: 5720
Sequence: 1 atctagcagcagcgagca.....gccttgcccaattagtattac 5720

Scoring table: IDENTITY_NUC
Gapop 10.0, Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 882724

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued Patents_NA.*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
5: /cgn2_6/ptodata/1/ina/PT05_COMB.seq.*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred: No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	5708	99.8	5720	US-09-442-100-1	Sequence 1, Appli
2	567.8	9.9	3155	US-09-442-100-7	Sequence 7, Appli
3	557	9.7	1961	US-09-509-902A-15	Sequence 15, Appli
4	536	9.4	1498	US-09-509-902A-6	Sequence 6, Appli
5	522.8	9.1	3213	US-09-442-100-5	Sequence 5, Appli
6	470.6	8.2	3984	US-09-442-100-3	Sequence 3, Appli
7	273.6	4.8	638	US-09-328-111-26	Sequence 26, Appli
8	261.2	4.6	2101	US-08-860-150-1	Sequence 1, Appli
9	261.2	4.6	2101	US-09-338-132-1	Sequence 1, Appli
10	254.8	4.5	2160	US-09-588-256-1	Sequence 1, Appli
11	232.6	4.1	3018	US-08-860-150-6	Sequence 6, Appli
12	232.6	4.1	3018	US-09-338-132-6	Sequence 6, Appli
13	216.6	3.8	1935	US-08-878-989-11	Sequence 11, Appli
14	216.6	3.8	1935	US-09-272-796-11	Sequence 11, Appli
15	164.6	2.9	678	US-09-328-111-66	Sequence 66, Appli
16	135.8	2.4	3182	US-08-484-044-11	Sequence 11, Appli
17	134.6	2.4	2726	US-08-423-699A-12	Sequence 12, Appli
18	134.6	2.4	2726	US-08-422-706B-12	Sequence 12, Appli
19	132.2	2.3	2511	US-08-422-699A-8	Sequence 8, Appli
20	132.2	2.3	2511	US-08-422-706B-8	Sequence 8, Appli
21	126.8	2.2	1599	US-09-256-465-1	Sequence 1, Appli
22	126.8	2.2	1599	US-09-167-322-3	Sequence 3, Appli
23	120.6	2.1	2549	US-09-467-082-3	Sequence 3, Appli
24	116.4	2.0	1732	US-09-430-564-1	Sequence 1, Appli
25	113.6	2.0	3489	US-08-728-323A-1	Sequence 1, Appli
26	113.6	2.0	3489	US-09-298-568-1	Sequence 1, Appli
27	113.6	2.0	32207	US-08-770-379-20	Sequence 20, Appli

c 28	113.6	2.0	32207	4	US-08-757-669A-20	Sequence 20, Appli
c 29	113.6	2.0	32207	4	US-09-230-371A-20	Sequence 20, Appli
c 30	112.8	2.0	4739	3	US-08-685-871-1	Sequence 1, Appli
31	110.8	1.9	4363	2	US-08-685-576-5	Sequence 5, Appli
32	110.2	1.9	2104	4	US-09-313-930-1	Sequence 1, Appli
33	110	1.9	2610	2	US-09-212-771-1	Sequence 1, Appli
34	110	1.9	2610	3	US-09-091-058-1	Sequence 1, Appli
35	109.2	1.9	5053	2	US-08-685-576-2	Sequence 2, Appli
36	107	1.9	1637	2	US-08-966-316-10	Sequence 10, Appli
37	106.2	1.9	1891	3	US-09-289-466-1	Sequence 1, Appli
38	104.6	1.8	1890	3	US-09-289-466-2	Sequence 2, Appli
39	104.6	1.8	1929	2	US-09-016-000-10	Sequence 10, Appli
40	103.8	1.8	2244	3	US-09-094-714A-48	Sequence 48, Appli
41	103.8	1.8	2245	4	US-09-225-749-24	Sequence 24, Appli
42	103.8	1.8	2599	6	5266464-1	Patent No. 5266464
43	102	1.8	477	4	US-09-135-994-1	Sequence 1, Appli
44	101.4	1.8	397	3	US-09-253-691-3	Sequence 3, Appli
45	99	1.7	2557	4	US-08-464-954A-1	Sequence 1, Appli

ALIGNMENTS

RESULT 1
US-09-442-100-1
; Sequence 1, Application US/09442100
; Patent No. 6359193
; GENERAL INFORMATION:
; APPLICANT: Xu, Tian
; APPLICANT: Tao, Wufan
; APPLICANT: Wang, Weiyl
; APPLICANT: Zhang, Sheng
; APPLICANT: Yu, Wan
; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
; NUMBER OF SEQUENCES: 16
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Pennie & Edmonds
; STREET: 1155 Avenue of the Americas
; CITY: New York
; STATE: New York
; COUNTRY: USA
; ZIP: 10036-2711
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patentin Release #1.0, Version #1.30
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/442.100
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/411.111
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Mistock, S. Leslie
; REGISTRATION NUMBER: 18,872
; REFERENCE/DOCKET NUMBER: 6523-003
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (212) 790-9090
; TELEFAX: (212) 869-9741/8864
; TELEX: 66141 PENNIE
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 5720 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: unknown
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1103..4402

US-09-442-100-1

Query Match 99.8%; Score 5708; DB 4; Length 5720;
Best Local Similarity 99.5%; Pred. No. 0;
Matches 5693; Conservative 26; Mismatches 1; Indels 0

Qy	1	ATCTAGCAGCGGAGCAACAAACACGAAATTAATTTTACTAAATTTAAAGCCAAACGC	60
Db	1	ATCTAGCAGCGGAGCAACAAACACGAAATTAATTTTACTAAATTTAAAGCCAAACGC	60
Qy	61	GCATCGGAAATGCCTGAAATGCGATTGAATGCACGCGAAAGTGATGGTTGCCAAGC	120
Db	61	GCATCGGAAATGCCTGAAATGCGATTGAATGCACGCGAAAGTGATGGTTGCCAAGC	120
Qy	121	GAGTGAATCAAGTGAAATACGTGCGCAATATACGCGAATTCGCTCAAAAGGCAAGGA	180
Db	121	GAGTGAATCAAGTGAAATACGTGCGCAATATACGCGAATTCGCTCAAAAGGCAAGGA	180
Qy	181	AAACGGAGAAAAAGAGAAAGCAATAAGTGCCTGTGTGGGAAACGCGAAAAAGGCGA	240
Db	181	AAACGGAGAAAAAGAGAAAGCAATAAGTGCCTGTGTGGGAAACGCGAAAAAGGCGA	240
Qy	241	GAACAAAGAGCGAAAAAGCGAGAAATTCGCTGGAAGCTGGAACGCGAAGAACGGA	300
Db	241	GAACAAAGAGCGAAAAAGCGAGAAATTCGCTGGAAGCTGGAACGCGAAGAACGGA	300
Qy	301	AGCTCCAAGTTGGCGCCATCGATTCGTGYGTAGGATCAATTAAGATTCCGAGTGGT	360
Db	301	AGCTCCAAGTTGGCGCCATCGATTCGTGYGTAGGATCAATTAAGATTCCGAGTGGT	360
Qy	361	GAATCGGCTCAAAATCAAAATTAATCAACTTAATTTGGTATTCAGATATTCAAATGGA	420
Db	361	GAATCGGCTCAAAATCAAAATTAATCAACTTAATTTGGTATTCAGATATTCAAATGGA	420
Qy	421	ATTCAATTCATCGCTGGAGCTTTTATTCGATCTGCCACTATTTTGGATTTGAAATGAT	480
Db	421	ATTCAATTCATCGCTGGAGCTTTTATTCGATCTGCCACTATTTTGGATTTGAAATGAT	480
Qy	481	GTGTCTCGGCTGGCGAGAAATCTGTATAAGCAGAGGAATAAATCGGAAGAACAA	540
Db	481	GTGTCTCGGCTGGCGAGAAATCTGTATAAGCAGAGGAATAAATCGGAAGAACAA	540
Qy	541	AATACAAATACAATGAAATGCGGGAGCAGTATTTACATGCGCAATCAATGCTGGATG	600
Db	541	AATACAAATACAATGAAATGCGGGAGCAGTATTTACATGCGCAATCAATGCTGGATG	600
Qy	601	CGNAAGGGGGTTCTCTTATATGCAATGTGAATGCGAATCGGAATCGGAATCGGA	660
Db	601	CGNAAGGGGGTTCTCTTATATGCAATGTGAATGCGAATCGGAATCGGAATCGGA	660
Qy	661	GTGGAAGAAATTCGCGCGGAGTGATAATATCCGACGCAACAAACGAGAGAGCCTAC	720
Db	661	GTGGAAGAAATTCGCGCGGAGTGATAATATCCGACGCAACAAACGAGAGAGCCTAC	720
Qy	721	ACGCGGAAAGAGCAGCGCAACAAATATCTTTATTGAGAGCAACAATATCAAGATC	780
Db	721	ACGCGGAAAGAGCAGCGCAACAAATATCTTTATTGAGAGCAACAATATCAAGATC	780
Qy	781	GAGATAATAAGCATCCTAAAACCGCGCTTAGTTCGTTATGCTCGCCACGATATA	840
Db	781	GAGATAATAAGCATCCTAAAACCGCGCTTAGTTCGTTATGCTCGCCACGATATA	840
Qy	841	GATATTCAAAGGCAAAAGGTGGTTCGGCATCGCCAGCAACAAAGTAAGCATCTATT	900
Db	841	GATATTCAAAGGCAAAAGGTGGTTCGGCATCGCCAGCAACAAAGTAAGCATCTATT	900
Qy	901	TCATACAAACCAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATA	960
Db	901	TCATACAAACCAACCAATTAATAATAATAATAATAATAATAATAATAATAATAATA	960
Qy	961	CAAAATCAATTCGCGCGCCGATGCGCCAGTGTGTGCGGTGTGTGTGTGTGTGTGTG	1020
Db	961	CAAAATCAATTCGCGCGCCGATGCGCCAGTGTGTGCGGTGTGTGTGTGTGTGTGTG	1020

QY 2101 CACCTCGGGGCTCGCGAGCCCCATAACTGTGTCTGCTGCGCGCGCGCCCTGGCGAA 2160
DB 2101 CACCTCGGGGCTCGCGAGCCCCATAACTGTGTGTGCTGCGCGCGCGCCCTGGCGAA 2160
QY 2161 GCCAACAACACAGTCTACAGCGCCAGGAGTACAGAGCCGATCATCATGAGTGTGAA 2220
DB 2161 GCCAACAACACAGTCTACAGCGCCAGGAGTACAGAGCCGATCATCATGAGTGTGAA 2220
QY 2221 GAGCAGCAGAGTCTCAAAAGCCGCTGCTCAAAACAGCAGTGGCGCGCAATGCCATCGAG 2280
DB 2221 GAGCAGCAGAGTCTCAAAAGCCGCTGCTCAAAACAGCAGTGGCGCGCAATGCCATCGAG 2280
QY 2281 TGCCTCGGCCAGCAATCACAGTCCAGTGTGCGCGCTCCACCTCTTACCTCAGAA 2340
DB 2281 TGCCTCGGCCAGCAATCACAGTCCAGTGTGCGCGCTCCACCTCTTACCTCAGAA 2340
QY 2341 GTCGCGGCGAGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2400
DB 2341 GTCGCGGCGAGTGTGACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2400
QY 2401 TCAGCAGCAGCAATCCAAACACCAACGCMACCAACACCGCCCTTGGTGGGTCTGAAYAG 2460
DB 2401 TCAGCAGCAGCAATCCAAACACCAACGCMACCAACACCGCCCTTGGTGGGTCTGAAYAG 2460
QY 2461 CAAGCCCAATTCGCTGGAGCCACCTCTCTATGCCAAGAGCATGAGGCGCAAGCGGCCAC 2520
DB 2461 CAAGCCCAATTCGCTGGAGCCACCTCTCTATGCCAAGAGCATGAGGCGCAAGCGGCCAC 2520
QY 2521 GGTGTARAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
DB 2521 GGTGTARAGCAGCAGCAACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGTGA 2580
QY 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2640
DB 2581 ACAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 2640
QY 2641 TCAGAGGAGCGGGATCAACGGGAGCGGGAACGGGATCAGCAGAGCTGGCCCAACGGAAA 2700
DB 2641 TCAGAGGAGCGGGATCAACGGGAGCGGGAACGGGATCAGCAGAGCTGGCCCAACGGAAA 2700
QY 2701 TCCTGGCGGAGAGTGTCTCGCGCGCGCGCCCTATCAGAGCAACACACACAGCAGGA 2760
DB 2701 TCCTGGCGGAGAGTGTCTCGCGCGCGCGCCCTATCAGAGCAACACACACAGCAGGA 2760
QY 2761 GATCAAAACCGCGAGCTGCAACAAACAAACATACAGATAGCAACAGCAACCTGGCGAC 2820
DB 2761 GATCAAAACCGCGAGCTGCAACAAACAAACATACAGATAGCAACAGCAACCTGGCGAC 2820
QY 2821 GACACACCCATTCGCGCTGCGCAATACAAATACAACTCTCTCAACACGGGCGCGCAATAG 2880
DB 2821 GACACACCCATTCGCGCTGCGCAATACAAATACAACTCTCTCAACACGGGCGCGCAATAG 2880
QY 2881 CTCGGGCGCAGCAGGATCCACGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2940
DB 2881 CTCGGGCGCAGCAGGATCCACGGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 2940
QY 2941 GATCAAGCAGCAGCTCGCCCATCCCGGAGCGCAAGAGATCTCAAGAGAGAGGAGGA 3000
DB 2941 GATCAAGCAGCAGCTCGCCCATCCCGGAGCGCAAGAGATCTCAAGAGAGAGGAGGA 3000
QY 3001 GCGCAAGGAGTTCGCGATCAGCAGTACTCTGCGCGCAAGCCTTCAAGTTCTTCATGGAGCA 3060
DB 3001 GCGCAAGGAGTTCGCGATCAGCAGTACTCTGCGCGCAAGCCTTCAAGTTCTTCATGGAGCA 3060
QY 3061 GCACATAGAGACGTGATCAAGTCTGATCGCCAGCGCAGGTATCGCAAGATCAGCTGGA 3120
DB 3061 GCACATAGAGACGTGATCAAGTCTGATCGCCAGCGCAGGTATCGCAAGATCAGCTGGA 3120
QY 3121 GAAGGAGATGCAAAAGTGGGACTCCCGATCAGACCCCAAAATCGAGATGAGGAAAATGCT 3180
DB 3121 GAAGGAGATGCAAAAGTGGGACTCCCGATCAGACCCCAAAATCGAGATGAGGAAAATGCT 3180
QY 3181 GAACCAAAAGGAGCAACTACATTCGATTGAAGCGCGCAAGATGGAGAGCATGTT 3240

DB 3181 GAACCAAAAGGAGCAACTACATTCGATTGAAGCGCGCAAGATGGACAGCATGTT 3240
QY 3241 CGTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAAAGCTGGTGGCAAAAT 3300
DB 3241 CGTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAAAGCTGGTGGCAAAAT 3300
QY 3301 CGATACCTCGAACCATTGTTATGCGATGAAACCCCTCGCGAAGCGAGCTTCTCAAGCG 3360
DB 3301 CGATACCTCGAACCATTGTTATGCGATGAAACCCCTCGCGAAGCGAGCTTCTCAAGCG 3360
QY 3361 GAATCAGTGGCACAGCTGAAGCGCGAGAGGATATCTTCGCGAAGCGCAAAATACTG 3420
DB 3361 GAATCAGTGGCACAGCTGAAGCGCGAGAGGATATCTTCGCGAAGCGCAAAATACTG 3420
QY 3421 GGTGTGAAGTGTACTACAGCTTCAGGACAAGGATATCTGTTGTGATGGACTA 3480
DB 3421 GGTGTGAAGTGTACTACAGCTTCAGGACAAGGATATCTGTTGTGATGGACTA 3480
QY 3481 CATACAGTGTGATCTGATCTGCTCATCAACTGGGCATTTTCGAGGAGGAAT 3540
DB 3481 CATACAGTGTGATCTGATCTGCTCATCAACTGGGCATTTTCGAGGAGGAAT 3540
QY 3541 GGCAGATTTCTACATCGCGAGGTCACTGCGCGCTGGACACGTTTCACAAAATGGGCTT 3600
DB 3541 GGCAGATTTCTACATCGCGAGGTCACTGCGCGCTGGACACGTTTCACAAAATGGGCTT 3600
QY 3601 CATTCACAGAGACATCAAGCCTGACAACATCTCATCTAGTAGGAGCGGACACATAAGCT 3660
DB 3601 CATTCACAGAGACATCAAGCCTGACAACATCTCATCTAGTAGGAGCGGACACATAAGCT 3660
QY 3661 CACCGACTTTGGCTGTGACGGGATTCGATGGAGCAACACTCGAAGTACTTACCAGGA 3720
DB 3661 CACCGACTTTGGCTGTGACGGGATTCGATGGAGCAACACTCGAAGTACTTACCAGGA 3720
QY 3721 GAACGGCAATCACTCGCGCAGGACTCGATGGAGCCCTGGGAGGAATCTCCGAGAACGG 3780
DB 3721 GAACGGCAATCACTCGCGCAGGACTCGATGGAGCCCTGGGAGGAATCTCCGAGAACGG 3780
QY 3781 ACCGAAGCCACCGTGTGAGAGCGGAGCGGATGCGGATCAACAAAGAGTCTTGGCCCA 3840
DB 3781 ACCGAAGCCACCGTGTGAGAGCGGAGCGGATGCGGATCAACAAAGAGTCTTGGCCCA 3840
QY 3841 CTCGCTGTGGGCAACCCGCAACTACATAGCTCCGAGGTGCTGGAGAGAGTGGGTACAC 3900
DB 3841 CTCGCTGTGGGCAACCCGCAACTACATAGCTCCGAGGTGCTGGAGAGAGTGGGTACAC 3900
QY 3901 GCAGCTGTGCGACTACTGGAGCGTGGGCTCATCTCTTACGAGATGCTGGTGGGTACGCC 3960
DB 3901 GCAGCTGTGCGACTACTGGAGCGTGGGCTCATCTCTTACGAGATGCTGGTGGGTACGCC 3960
QY 3961 GCCCTTCTGGCCCAACAGTCCGCTGGAACGCAACAAAGGTCTCACTGGGAGAAAC 4020
DB 3961 GCCCTTCTGGCCCAACAGTCCGCTGGAACGCAACAAAGGTCTCACTGGGAGAAAC 4020
QY 4021 SCTGCATATTCGCGCGCAGCGGAGTATTCGCGGAGGCTACGAGCTTGTATAGAGGCT 4080
DB 4021 SCTGCATATTCGCGCGCAGCGGAGTATTCGCGGAGGCTACGAGCTTGTATAGAGGCT 4080
QY 4081 CTGTCGCTGGCTGACAGCGGCTGGCAAGAGCGTGGACAGGCTCAAGAGCCACGACTT 4140
DB 4081 CTGTCGCTGGCTGACAGCGGCTGGCAAGAGCGTGGACAGGCTCAAGAGCCACGACTT 4140
QY 4141 CTTCAAGGSCATCGACTTTCGCGGACATCGGAGCAGAAAGCGCCCTACATACCCGAAT 4200
DB 4141 CTTCAAGGSCATCGACTTTCGCGGACATCGGAGCAGAAAGCGCCCTACATACCCGAAT 4200
QY 4201 CAAGCACCACRACGACACATCCAACTTTGATCCCGTGTATCCGAGAGCTCGCTCGAA 4260
DB 4201 CAAGCACCACRACGACACATCCAACTTTGATCCCGTGTATCCGAGAGAGCTCGCTCGAA 4260
QY 4261 TGACTCCACCATGAGCAGCGGATGATGTCGACCAAGATGACCGCACYTTCCACGGCTT 4320
DB 4261 TGACTCCACCATGAGCAGCGGATGATGTCGACCAAGATGACCGCACYTTCCACGGCTT 4320

Fri Jan 17 11:17:10 2003

Db 4261 TGACTCCACCATGAGCGGGCGATGATGTCGACCAAGATGACCGACATTTCCACGGCTT 4320
Qy 4321 TTTGGAATTTACCTTCCTGCTGCTTCTTCGACGACAGCGCGCGGATATGACGGACGA 4380
Db 4321 TTTGGAATTTACCTTCCTGCTGCTTCTTCGACGACAGCGCGCGGATATGACGGACGA 4380
Qy 4381 TCAGGCGCGGTTTACGTCCTGAAATGATGCTCTCCATGTCGCCAACACACACACCCCGC 4440
Db 4381 TCAGGCGCGGTTTACGTCCTGAAATGATGCTCTCCATGTCGCCAACACACACACCCCGC 4440
Qy 4441 CCCCAGATCATGTTAGTCAAAATAGTCACAAAAGGGGATAGAAACCATTTAGTGGGCTT 4500
Db 4441 CCCCAGATCATGTTAGTCAAAATAGTCACAAAAGGGGATAGAAACCATTTAGTGGGCTT 4500
Qy 4501 GCATTTGTAAGGAGCGGTGCTATAGAAATGAAACTATCTATATACATTTATAAATATA 4560
Db 4501 GCATTTGTAAGGAGCGGTGCTATAGAAATGAAACTATCTATATACATTTATAAATATA 4560
Qy 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATTAATATACATATATTTGAT 4620
Db 4561 GGAGACAGTAGAGCGGGAGCTAGCTATATACATACAAATTAATATACATATATTTGAT 4620
Qy 4621 ATATATATATATATGCGGTAGGAGCTAGCTGAATTAATATAAAGCGGAGCTAG 4680
Db 4621 ATATATATATATATGCGGTAGGAGCTAGCTGAATTAATATAAAGCGGAGCTAG 4680
Qy 4681 AGATGAAACGAGAGCGGAGCTAGGACCTTCGACCTTTAACTGAACATAGTATCTT 4740
Db 4681 AGATGAAACGAGAGCGGAGCTAGGACCTTCGACCTTTAACTGAACATAGTATCTT 4740
Qy 4741 GTGCACTACTCTCACAAATATATATTTTAAATTTAGAAATTCAAAAGGGGCCA 4800
Db 4741 GTGCACTACTCTCACAAATATATATTTTAAATTTAGAAATTCAAAAGGGGCCA 4800
Qy 4801 ACTGGAATCGAACCTTTCTGCTGCTCAAGCAAGCAAGCAAGCAAGCAAGCAAGCC 4860
Db 4801 ACTGGAATCGAACCTTTCTGCTGCTCAAGCAAGCAAGCAAGCAAGCAAGCC 4860
Qy 4861 TTAACATAAAGGAGCGGAATTTACCAACCACTTCACTCTCTCTCTCTCTCTCTCT 4920
Db 4861 TTAACATAAAGGAGCGGAATTTACCAACCACTTCACTCTCTCTCTCTCTCTCTCT 4920
Qy 4921 GATCGTGGCGGATTCGAATCAGCAGCGCTGTTGCATCCGGCCATCCCATTTACTTCC 4980
Db 4921 GATCGTGGCGGATTCGAATCAGCAGCGCTGTTGCATCCGGCCATCCCATTTACTTCC 4980
Qy 4981 CATTCAGAAATGAGATTCGAGGTGTCGATGAGACGAGCGAGACCAAGTCGCAC 5040
Db 4981 CATTCAGAAATGAGATTCGAGGTGTCGATGAGACGAGCGAGACCAAGTCGCAC 5040
Qy 5041 GCGAGCGATATAGCGGCTTTAATAGCCTTAATCTAAATCTAACTGGGAGAACAGGACC 5100
Db 5041 GCGAGCGATATAGCGGCTTTAATAGCCTTAATCTAAATCTAACTGGGAGAACAGGACC 5100
Qy 5101 YATGATGCTCTGCTATCCAAATTCGCTATCATCTGCTTCACTGCTGCTGCTGCTGCT 5160
Db 5101 YATGATGCTCTGCTATCCAAATTCGCTATCATCTGCTTCACTGCTGCTGCTGCTGCT 5160
Qy 5161 CCCCCCTCCCATCCAAAGCAACAACTTACAGCTAGCTATGTGAAAGCTAGCAAT 5220
Db 5161 CCCCCCTCCCATCCAAAGCAACAACTTACAGCTAGCTATGTGAAAGCTAGCAAT 5220
Qy 5221 GTTAGACCAACTTGTGAATGCCAAATGAAATGTTTACGCCACGAGGAAACGGGG 5280
Db 5221 GTTAGACCAACTTGTGAATGCCAAATGAAATGTTTACGCCACGAGGAAACGGGG 5280
Qy 5281 GAAATTTCAACTTATTTCTGTAGCAACAGGAAAGAAAGAAAGAAAGAAAGAAAGAG 5340
Db 5281 GAAATTTCAACTTATTTCTGTAGCAACAGGAAAGAAAGAAAGAAAGAAAGAAAGAG 5340
Qy 5341 AAACAGTACGAGAAATTTGTAATCTTCTTATGTAATGTAATGTAATGTAATGTAATG 5400
Db 5341 AAACAGTACGAGAAATTTGTAATCTTCTTATGTAATGTAATGTAATGTAATGTAATG 5400

Qy 5401 AATCTATGCTAGAGCTTGTGTAGCGCCCTTAAGATGTTTTTTTAGTTTATAGACCCTAACCG 5460
Db 5401 AATCTATGCTAGAGCTTGTGTAGCGCCCTTAAGATGTTTTTTTAGTTTATAGACCCTAACCG 5460
Qy 5461 TAATCTAGCTTTAATTCCTAACACATAAGCGAGAGTACAGTACATGTTTGTGTTGTC 5520
Db 5461 TAATCTAGCTTTAATTCCTAACACATAAGCGAGAGTACAGTACATGTTTGTGTTGTC 5520
Qy 5521 GTAGTTCGTTGGAAATGCTTAACGGAAGAGATTTGTTTTCTCTTTAATTAGCTTCA 5580
Db 5521 GTAGTTCGTTGGAAATGCTTAACGGAAGAGATTTGTTTTCTCTTTAATTAGCTTCA 5580
Qy 5581 GTTTCATGTCGCTGTTTTTATATGACTTATATATAGTCCATCTGAATATTCGTGGA 5640
Db 5581 GTTTCATGTCGCTGTTTTTATATGACTTATATATAGTCCATCTGAATATTCGTGGA 5640
Qy 5641 TGGAGCCTATTTAAATGTGAGATCGAGCTAAATTAAGGAAATACAAACAACCTCTGT 5700
Db 5641 TGGAGCCTATTTAAATGTGAGATCGAGCTAAATTAAGGAAATACAAACAACCTCTGT 5700
Qy 5701 GCCTTGGCCCAATTAGTTTAC 5720
Db 5701 GCCTTGGCCCAATTAGTTTAC 5720

RESULT 2
US-09-442-100-7
Sequence 7, Application US/09442100
Patent No. 6359193
GENERAL INFORMATION:
APPLICANT: Xu, Tian
APPLICANT: Tao, Wufan
APPLICANT: Wang, Welyi
APPLICANT: Zhang, Sheng
APPLICANT: Yu, Wan
TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
TITLE OF INVENTION: GENES AND METHODS BASED THEREON
NUMBER OF SEQUENCES: 16
CORRESPONDENCE ADDRESS:
ADDRESSEE: Pennie & Edmonds
STREET: 1155 Avenue of the Americas
CITY: New York
STATE: New York
COUNTRY: USA
ZIP: 10036-2711
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/442,100
FILING DATE:
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 08/411,111
FILING DATE:
ATTORNEY/AGENT INFORMATION:
NAME: Mirock, S. Leslie
REGISTRATION NUMBER: 18,872
REFERENCE/DOCKET NUMBER: 6523-003
TELECOMMUNICATION INFORMATION:
TELEPHONE: (212) 790-9090
TELEFAX: (212) 869-9741/8864
TELEX: 66141 PENNIE
INFORMATION FOR SEQ ID NO: 7:
SEQUENCE CHARACTERISTICS:
LENGTH: 3155 base pairs
TYPE: nucleic acid
STRANDEDNESS: double
TOPOLOGY: unknown
MOLECULE TYPE: cDNA

FEATURE: CDS
NAME/KEY: 1..2943
LOCATION: 1..2943
US-09-442-100-7

Query Match 9.9%; Score 567.8; DB 4; Length 3155;
Best Local Similarity 64.1%; Pred. No. 6.8e-126;
Matches 978; Conservative 4; Mismatches 491; Indels 48; Gaps 7;

QY 2869 GGGCGGATAGCTCGGGCGGAGCAAGATCCACCGGACACCGCTCTCTCTCGTGCAC 2928
DB 1329 GGGCGGATAGCTCGGGCGGAGTACAGAGCCACAAAGGTGGGAGGAGACAAAGCTGG 1388
QY 2929 CAGCTCCAAAGAGATCAAGCAGCGCTCGCCCATCCCGGAGCGCAAGAGATCTCCAAGGA 2988
DB 1389 CAGAGACAAAGCAGATTCAGACCTCCCGGCTGCTCGGCAAGATAGC-----1440
QY 2989 GAAGGAGGAGGCGCAAGAGATTCGCGCATCAGGCGAGTACTCGCCCAAGCCCTTCAAGTT 3048
DB 1441 -AGAGATGAAGAGAGAGAGTCTCGCATCAAGAGTACTCGCCCTTATGCGCTTCAAAAT 1499
QY 3049 CTTATGGAGCAGCATAGAGAGCTGATCAAGTCTGATCGCCAGCGCACGTATCGCAA 3108
DB 1500 CTTATGGAGCAGCATAGAGAGTCTCAAAAGCTTCAAAAGCTTCAAGAGGTTCAGCCGGAG 1559
QY 3109 GAATCAGCTGGAGAGGAGTGCACAAAGTGGGACTCGCCGATCAGACCCCAATTCGAGAT 3168
DB 1560 GCTACAGCTGGAGCAGAAATGCCAAAGCTGGGCTCTGTAGGCGGAGCAGGAGCAGAT 1619
QY 3169 GAGGAAATGCTGAACAAAGAGAGCAACTACATTCGATTGAAGCGCGCCCAAGATGA 3228
DB 1620 GAGGAAATGCTTACCAAGAGAGTCTAACTACAAAGCGGTGAAGAGGCGCAAGATGA 1679
QY 3229 CAAGCAGTCTCGTCAAACTGAGCCCAATGGAGTGGGTGATTTGGCGAGGTACGCT 3288
DB 1680 CAAGTCCATGTTGTGAAATCAAGACTCTAGGCATCGGTGCTTTGGGGAAGTGTGCCT 1739
QY 3289 GGTGAGCAAAATCGATACCTCGAACCATTTGTATGGGATGAAACCTCGGAAAGCGGA 3348
DB 1740 CGCTTGTAGCTGA---CACTCAGCTGTGTACGCCATGAAGCTCTCAGGAGNAGGA 1796
QY 3349 CGTTCAGGCGAATCAGTGGGACACGTCGAGGCGGAGGAGATATCTCGCGGAAGC 3408
DB 1797 TGTCTGAAACCGGAATCAAGTGGCCATGTCAAGGCTGAGAGGAGATCTTGCTGAAGC 1856
QY 3409 CGACATACTGGTGGTGAAGTGTGACTACAGCTTCCAGGACAGGATTAATCTGACTT 3468
DB 1857 AGCAATAGTGGTGGTCAAACTCTACTCTCTTCCAGGACAGGACAGGCTGTACTT 1916
QY 3469 TGTGATGAGTACATACAGGCTGGTGTATGATGTGCTGCTCATCAAACTGGGCTTTT 3528
DB 1917 TGTGATGAGTACATACAGGCGGGATATGATGAGCTGTCTGATCAGGATGGAGTCTT 1976
QY 3529 CGAGGAGAACTGGCGAGATTTACATCGCGAGGTACCTGGCGGCTGGACAGCGTTCA 3588
DB 1977 CCGTGAACCTGGCGCGCTTCTACATTCGAGAGTTGACCCCTGGCCATGAAAGTGTCCA 2036
QY 3589 CAAATGGGCTTCATTCAGAGACATCAAGCTGACACATACATCTCATGATAGGACGG 3648
DB 2037 CAAGATGGGCTTTATCCACGGGACATCAAGCTTGACACATACATCTGACGCTGGATGG 2096
QY 3649 ACATATAAGCTCACCGACTTTGGCTGTGCGAGGATCCGATGGACGCAACATCGAA 3708
DB 2097 TCATATTAAGCTGACAGATTTGGCTCTGCTGCTGATGAGTTCAGTGGACATCAATTCAA 2156
QY 3709 GTACTACGAGGAAACCGCAATCTACTCGCCGAGGACTCGATGGAGCCCTGGAGGAAT- 3767
DB 2157 GTACTACC---AGAAAGGAACCAATGAGACAGGACATGGAGCCCGGTGACCTCTG 2213
QY 3768 -----ACTCCGAGAACGGACCGAGCCGACCGCTGCTGGAGGGGAGCGAT 3813
DB 2214 GGACGATGTTCCAACTGTCGTGTGAGACAGGTTAAAGACCCCTGGAGCAGAGGCGCA 2273

QY 3814 GCGGATACCAAAGAGTCTTGCCACATCTGCTGGTGGCAGCCCGCAACTACATAGCTCC 3873
DB 2274 GAAGCAGCAGAGAGTCTGCGACATTTCTTGTGGGACACCAATATACATCGCTCC 2333
QY 3874 CGAGTGTGGAGAGAGTGGGTACACGAGCTGTGCGACTACTGGAGCGTGGGCGTCTAT 3933
DB 2334 GGAGTGTCTTCTCGCAAGAGGTACACGAGCTCTGTACTGGTGGAGCGTGGTGTGAT 2393
QY 3934 CCTTATAGATGCTGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3993
DB 2394 TCTTCTTGTAGATGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 2453
QY 3994 ACAAAGGTTCATCACTGGGAGAAACCTGCATATTCGCGCGAGCGGAGTTATCCGG 4053
DB 2454 GCTGAGGTGATCACTGGGAGAGCAGCTGCATATCCCTAGCGAGTGGCTCAGCGC 2513
QY 4054 CGAGGCTACGAGCTTGTATAAGGAGGCTGTGCGTGGTGGTGGTGGTGGTGGTGGTGG 4110
DB 2514 TGAGGCGCGAGACCTCATCAGAAAGCTGTGCTGGCGGCTGACTGCGGCTGGCAGGA 2573
QY 4111 GAGCGTGGAGAGGTCAAGAGCAGCAGCTTCTTCAAGGCGATCGACTTTGC---GGACAT 4167
DB 2574 TGGGCGAGTACCTCAAGGACACCCGCTTCTTCAACACCATCGACTTTTCCCGTGACAT 2633
QY 4168 GCGGAGCAGAAAGCGCCCTACATACCGGAAATCAAGCACCCACGAGACACATCCAACTT 4227
DB 2634 CGAAAGCAGCTGCACCTAGTCTCCACCATCAGCCACCCCATGGACACCTCCATTT 2693
QY 4228 TGATCCGCTGGATTCGGGAGAGCTGGCTGCAATGATCTCCACATGACGAGCGGATGA 4287
DB 2694 TGACCCGCTGGATGAAGAAAGCCCTGGCAGGAGGCGGAGAGAGGCGCAAGGCGCT 2753
QY 4288 TGTGACACAGATGACCGGACACTTTC-----CAGCGCTTTTGAATTTACCTT 4335
DB 2754 GGACAGCTGGGCTTCCCGCAGCAGCAAGCATCCAGAGCAGCGCTTCTATGAGTTACCTT 2813
QY 4336 CCGTGTCTTCTCGACGACAA 4356
DB 2814 CCGAGGTTCTCGATGACAA 2834

RESULT 3

US-09-509-902A-15
; Sequence 15, Application US/09509902A
; Patent No. 6387676
; GENERAL INFORMATION:

; APPLICANT: Virca, Duke

; APPLICANT: Bird, Timothy A.

; APPLICANT: Anderson, Dirk M.

; APPLICANT: Marken, John S.

; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions

; FILE REFERENCE: 2877-US

; CURRENT APPLICATION NUMBER: US/09/509,902A

; CURRENT FILING DATE: 1999-08-03

; NUMBER OF SEQ ID NOS: 16

; SOFTWARE: PatentIn Ver. 2.0

; SEQ ID NO 15

; LENGTH: 1961

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-509-902A-15

Query Match 9.7%; Score 557; DB 4; Length 1961;

Best Local Similarity 64.1%; Pred. No. 2e-123;

Matches 963; Conservative 4; Mismatches 494; Indels 42; Gaps 7;

QY 2884 GGGCGGCAACAGGATCCACCGGACACCGCTCTCTGTCGAGCTGCAAGAGAT 2943
DB 158 GGGCGGCAACAGGATCCACCGGACACCGCTCTCTGTCGAGCTGCAAGAGAT 217
QY 2944 CAAGCAGCGCTCGCCCATCCCGGAGCGCAAGAGATCTTCAAGGAGAGGAGGAGCG 3003
DB 218 GATTACAGACCTCTCCGTTCCCGTCCGCAAAACAGC-----AGAGACCAAGAGAA 268

QY 3004 CAAGGAGTTCGGATCAGGAGTACTCGCCGCAAGCCCTTCAAGTCTTCTCATGGAGACGA 3063
 Db 269 GAGAGAGTACGAGTCAAGAGTACTCGCCATACGCCCTTTAAGTCTTCTCATGGAGACGA 328
 QY 3064 CATAGAGACGATGATCAAGTCTGATCCGAGCGACGATATCGCAAGATCAGTGGAGAA 3123
 Db 329 CGTGAGAGTGTATCAAAACCTACACAGCAGAGGTTAAACCGAGGCTGCGAGCTGGAGCA 388
 QY 3124 GGAGATGCAAAAGTGGAGTGGCCGATCAGACCCCAATCGAGATGAGGAAATGCTGAA 3183
 Db 389 AGAAATGGCCAAAGCTGACTCTGTGAAGTGAAGAGGACGATCGGAAGATCTCTA 448
 QY 3184 CCAAAAGGAGCAACTACATTCGATGAAGCGCCGCAAGATGGAAGAGCATGTCGT 3243
 Db 449 CCAGAAAGAGTCTAATTAACAACAGGTTAAAGAGGCGCAAGATGGAACATCTATGTTCT 508
 QY 3244 CAAACTGAACCCATTTGGAGTGGGTGCTATTTGGCGAGGTAAACGCTGGTGAAGAAATCGA 3303
 Db 509 CAAGATCAAAACCCCTGGGATCGTGCCTTTGGAGAGTGTGCCCTTGTGAAGTGA 568
 QY 3304 TACTCTGAACCAATTTGTATGCGATGAACACCTCGCGAAGCGGAGCTTCTCAAGCGGAA 3363
 Db 569 ---CACTACAGCCCTGTAGCCATGAAGACCCCTAAGGAAAGAGATGCTCTCAACCGGA 625
 QY 3364 TCAGGTGGCAGACGTCGAAGCGGAGAGGATATCTCTCGCGGAGCGGACAAATACCTGGT 3423
 Db 626 TCAGGTGGCCAGCTCAAGGCGGAGAGGACATCTCGCCGAGGAGCAATGAGTGGT 685
 QY 3424 GGTGAAGTGTACTACAGCTCCAGGACAAGGATATCTGTACTTTGTGATGGACTACAT 3483
 Db 686 GGTCAAACTCTACTACTCTCTCCAGACAAAGACAGCTGTACTTTGTGAGGACTACAT 745
 QY 3484 ACCAGTGTGATGTAGTGCCTGCTCATCAAACTGGGCAATTTTCGAGGAGCACTGGC 3543
 Db 746 CCTGTGTGGGACATCATGAGCTGCTGATCGGATGGAGTCTTCCCTGAGCACTGGC 805
 QY 3544 CAGATTCTACATCGCGAGGTCACTCGCGGTGACAGCGCTTCAAAATTTGGCTTCAT 3603
 Db 806 CCGGTTCTACATCGAGAGTCACTTTGGCCATTTGAGAGTGTCCACAGATGGGCTTCAT 865
 QY 3604 TCACAGACATCAAGCTGTGACAACTACTCATCTAGTAGGAGCGACACATAAAGCTCAC 3663
 Db 866 CCACCGAGACATCAAGCTGTATCAATTTGATAGATCTGGATGTCACATTAACCTCAC 925
 QY 3664 CCACTTTGGCTGTGACGCGATTCGATGAGCGGCTTGGAGGAGTACTCCGA--- 3774
 Db 986 GAGCCATGTACAGACAGGACAGCATGGAGCCAGCGACCTCTGGATGATGTGCTAAGT 1045
 QY 3775 ---GAAACGAGGAGCCCGCTGTGGAGAGCGGAGGATCGCGATCAACCAAGAGT 3831
 Db 1046 TCGGTGTGGGACAGGCTGAAGACCTTAGAGAGAGGCGCGGAAGCAGCAGAGGTG 1105
 QY 3832 CCGGCGCACTCGCTGTGGTGGGACCCCGAACTACATAGCTCCCGAGTGTGGAGAGAG 3891
 Db 1106 CCGGCACTTCACTGTGGGAGTCCAACTACATCGACCCAGGAGTGTCTCCGCA 1165
 QY 3892 TGGGTACAGGAGTGTGGAGTACTGGAGCGTGGCGTCTATCTTAYAGATGCTGGT 3951
 Db 1166 AGGTACACTCACTCTGTGACTGTGGAGTGTGGAGTGTGATCTCTCGAGATGCTGT 1225
 QY 3952 GGGTCAGCGCCCTTTCTGGCCAAAGTCCGCTGGAAACCAACAAAGGTCATCAACTG 4011
 Db 1226 GGGCAGCCGCCCTTTTGGCACTTCTCCCAAGAAACCCAGCTGAAGGTGATCAACTG 1285
 QY 4012 GGAGAAACCTGCATATTTCCGCGGAGCGGAGTATCCCGGAGGCTACGAGCTGTAT 4071
 Db 1286 GGAGAACACCTCCATATCCAGCCAGGTTGAAGCTGAGCCCTGAGCCGAGGACCTCAT 1345

QY 4072 AAGGAGGCTCTGTGCTGGCTTGACAAGCGGCTGGCAAGAGCGTGGACGAGG---TCAA 4128
 Db 1346 CACCAAGCTGTGCTGCTCGCGACACACCGCTGGGGGGAATGGCGCGGATGACCTGAA 1405
 QY 4129 GAGCCACGAGTCTTCTCAAGGGGATCGACTTTGC---GACATCGCGGAAGCAGAAAGCC 4185
 Db 1406 GGCACCCCTCTCTCAGCGCATGACTTCTCCAGTGCATCCGGAAGCATCCAGGCC 1465
 QY 4186 CTACATACGGGAATCAAGCAACCCCGACACATCCAACTTTGATCCCGTGGATCCGGA 4245
 Db 1466 CTACGTTCCACCATCAGCCACCCCATGACACCTCGAATTTCCAGCCCGTAGATGAGA 1525
 QY 4246 GAAGCTGGCTGATGACTCCACCATCAGCAGCGGATGATGTC-----GA 4293
 Db 1526 AGCCCTTGGACGATGCGCAGGAGGTAGCAACAGCCCTGGGACACACTCACTCGCC 1585
 QY 4294 CCAGATACCGCGACCTTCCAGCGCTTTTTCGAATTTTACCTTCGCTGCTTCTTCGAGCA 4353
 Db 1586 CAATAACAAGCATCTCGAGCAGCATTTTACGATTTACCTTCCGAAGGTTCTTTTGATGA 1645
 QY 4354 CAA 4356
 Db 1646 CAA 1648

RESULT 4
 US-09-509-902A-6
 ; Sequence 6, Application US/09509902A
 ; Patent No. 6387676
 ; GENERAL INFORMATION:
 ; APPLICANT: Virca, Duke
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Anderson, Dirk M.
 ; APPLICANT: Marken, John S.
 ; TITLE OF INVENTION: Human cDNAs Encoding Polypeptides Having Kinase Functions
 ; FILE REFERENCE: 2877-US
 ; CURRENT APPLICATION NUMBER: US/09/509,902A
 ; CURRENT FILING DATE: 1999-08-03
 ; NUMBER OF SEQ ID NOS: 16
 ; SOFTWARE: PatentIn Ver. 2.0
 ; SEQ ID NO 6
 ; LENGTH: 1498
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-09-509-902A-6

Query Match 9.4%; Score 536; DB 4; Length 1498;
 Best Local Similarity 65.5%; Pred. No. 1.8e-118;
 Matches 884; Conservative 3; Mismatches 433; Indels 30; Gaps 6;

QY 2884 GGGCGGACGACGAGTCCACCGGACACCGCTCTCTGTCGACCGCTCAAGAGAT 2943
 Db 160 GGGCGGACGACGAGTCCAAAGCGCCGCAAGGCGGACAAAGCGGAAAGGATAAAGCA 219
 QY 2944 CAAGCAGCGCTCGCCATCCCGGAGCGCAAGAGATCTTCAAGAGAGAGGAGGAGCG 3003
 Db 220 GATTGAGACCTCTCCGCTTCCGCTCGGCAAAACAGC-----AGAGACGAGAGAA 270
 QY 3004 CAAGGATGTCGACGAGTCTGCGCGCAAGCCCTCAAGTCTTCATGAGGACGCA 3063
 Db 271 GAGAGAGTCAAGCATCAAGAGTCTGCGCATACCCCTTAAGTCTTCATGAGGACGCA 330
 QY 3064 CATAGAGAGTGTATCAAGTGTGCGCAGCGCAGCTATCGCAAGATCAAGTGGAGAA 3123
 Db 331 CGTGAGAGTGTATCAAAACCTTACAGCAAGGTTAAACCGAGGCTGCACTGGAGCA 390
 QY 3124 GGAGATGCAAAAGTGGGAGTGGCCGATGCGCAGCGCAGCTATCGCAAGATCAAGTGGAGAA 3183
 Db 391 AGAATGCCAAAGTGGAGTGTGAGCTGTGAAGCTGTGAGCAGGAGATGGGAAGATCTCTA 450
 QY 3184 CCAAAAGGAGGACACTACATTTGATTCGATTTGAAGCGCCCAAGATGCAAGAGCATGTCGT 3243
 Db 451 CCAGAAAGAGTCTAATTACAACAGGTTAAAGAGGCGCAAGATGACCAAGTCTATGTTGT 510

QY	3244	CAAACTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTAAAGCTGTGTGACGAAAATTCGA	3303
Db	511	CAAGATCAAAACCCCTGGGGATCGGTGCCTTTTGGAGAACTGTGCCTTGTCTTAAGTGGGA	570
QY	3304	TACCTCGAAACCATTTGATGCGCATGAAACCCCTCGGAAAGCGGACGTTCTCAAGCGGAA	3363
Db	571	---CACTACGCCCCTGACGCCATGAGACCCTAGGAAAGAGATGTCCTGAACCGGAA	627
QY	3364	TCAGGTGGCAGACGTGAAGCGCGAGAGGATATCTCCGCGAAAGCCGACAAATAACTGGGT	3423
Db	628	TCAGGTGGCCCACTCAAGCGCGAGAGGACATCTCGCCGAGGACACAAATGATGGGT	687
QY	3424	GGTGAAGTTGTACTACAGCTTCCAGGACAGGATAACTGTACTTTGTGTGATGGACTACAT	3483
Db	688	GGTCAAACTCTACTCTCTCCCAAGACAAGACAGCCTGTACTTTGTGTGATGGACTACAT	747
QY	3484	ACCAGTGGTGTATCTGATCTCGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGAACCTGGC	3543
Db	748	CCCTGCTGGGACATGATGAGCCTGCTGATCCGGATGGAGCTTCTCCCTGAGCACCCTGGC	807
QY	3544	CAGATTCTACATCGCGGAGTCACTGCGCCGCTGGACAGCGTTTCACAAAATGGGCTTCAT	3603
Db	808	CCGGTTCTACATCGCAGAGCTCACTTTGGCCATTGAGAGTGTCCACAAGATGGGCTTCAT	867
QY	3604	TCACAGAGACATCAAGCCTTGACAACATACTCATCATAGGACGAGCACACATAAAGCTCAC	3663
Db	868	CCACCAGACATCAAGCCTGATAACATTTGATAGATCTGGATGGTCACATTAACTCAC	927
QY	3664	CGACTTTGGCCTGTGCACGGGATTCGGATGGAGCGACAACTCGAAGTACTTACACAGA---	3720
Db	928	AGATTTGGGCCTCTGCACCTGGGTTTCAGGTGGACTCACAAATTCCAAATATTACCAGAAAGG	987
QY	3721	---GAACGGCAATCACTCGCGCCAGAGCTCGATGGAGCCCTGGGAGGA-----ATACTC	3771
Db	988	GAGCCATGTCAGACAGGACAGCATGGAGCCACGACCTCTGGGATGATGTGTCTAACTG	1047
QY	3772	CGAAGACGAGCAGAACGCCACCGTCTCGAGAGGCGCGGATGCGGATCACCAAGAGT	3831
Db	1048	TCGGTGTGGGACAGGCTGAAGACCCCTAGAGCAGAGGCGCGGAAGCAGACCAGAGGTG	1107
QY	3832	CTTGGCCCACTCGCTGTGGGCAACCCGAACTACATAGCTCCGAGGTGCTGGAGAGAG	3891
Db	1108	CTTGGCACAATTCATGTGGGGACTTCAAACATACATCGACCCGAGGTCTCTCCCGCAA	1167
QY	3892	TGGGTACACGACGTGTGCGACTACTTGGAGCGTGGCGCTATCTCTTAYGAGATGCTGGT	3951
Db	1168	AGGGTACACTCAACTCTGTGACTGTGTGGAGTGTGGAGTGATTCTCTTCGAGATGCTGGT	1227
QY	3952	GGGTGAGCGCCCTTTCGCGCAACAGTCCGCTGGAACGCAACAAGAGTCATCAACTG	4011
Db	1228	GGGGCAGCGCCCTTTTGGCACCTTACTTCCACAGAAACCCAGCTGAAGTGTGATCAACTG	1287
QY	4012	GGAGAAACACTGCAATATTCGCGCGCAGCGCGAGTTATTCGCGGAGGCTACGCACTTGAT	4071
Db	1288	GGAGAACAGCTCCACATTTCCAGCCCAAGTGAAGCTGAGCCCTCGAGCCAGGACCTCAT	1347
QY	4072	AAGGAGCTCTGTGCTGGCTGACAGCGGCTGGCGAAGAGCTGGACACAGG---TCAA	4128
Db	1348	CAGCAAGCTGTGCTGCTCGCAGACACCGCTCGGGCGGAATGGGCGCATGACCTGAA	1407
QY	4129	GAGCCACGACTTCTTCAAGGGCATCGACTTTTGC---GGACATCGGGAAGCAGAAAGCGCC	4185
Db	1408	GGCCACCCCTCTCTCAGCGCCATTGACTTCTCCAGTGACATCCGAAGCATCCAGGCC	1467
QY	4186	CTACATACCGGAAATCAAGCACCCACCGGA	4215
Db	1468	CTACGTTCCACCATCAGCCACCCCATGGA	1497

; Patent No. 6359193
 ; GENERAL INFORMATION:
 ; APPLICANT: Xu, Tian
 ; APPLICANT: Tao, Wufan
 ; APPLICANT: Wang, Weiyl
 ; APPLICANT: Zhang, Sheng
 ; APPLICANT: Yu, Wan
 ; TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
 ; TITLE OF INVENTION: GENES AND METHODS BASED THEREON
 ; NUMBER OF SEQUENCES: 16
 ; CORRESPONDENCE ADDRESS:
 ; ADDRESSEE: Pennie & Edmonds
 ; STREET: 1155 Avenue of the Americas
 ; CITY: New York
 ; STATE: New York
 ; COUNTRY: USA
 ; ZIP: 10036-2711
 ; COMPUTER READABLE FORM:
 ; MEDIUM TYPE: Floppy disk
 ; COMPUTER: IBM PC compatible
 ; OPERATING SYSTEM: PC-DOS/MS-DOS
 ; SOFTWARE: Patent In Release #1.0, Version #1.30
 ; CURRENT APPLICATION DATA:
 ; APPLICATION NUMBER: US/09/442,100
 ; FILING DATE:
 ; CLASSIFICATION:
 ; PRIOR APPLICATION DATA:
 ; APPLICATION NUMBER: 08/411,111
 ; FILING DATE:
 ; ATTORNEY/AGENT INFORMATION:
 ; NAME: Mirock, S. Leslie
 ; REGISTRATION NUMBER: 18,872
 ; REFERENCE/DOCKET NUMBER: 6523-003
 ; TELECOMMUNICATION INFORMATION:
 ; TELEPHONE: (212) 790-9090
 ; TELEX: (212) 869-9741/8864
 ; TELEFAX: 66141 PENNIE
 ; INFORMATION FOR SEQ ID NO: 5:
 ; SEQUENCE CHARACTERISTICS:
 ; LENGTH: 3213 base pairs
 ; TYPE: nucleic acid
 ; STRANDEDNESS: double
 ; TOPOLOGY: unknown
 ; MOLECULE TYPE: CDNA
 ; FEATURE:
 ; NAME/KEY: CDS
 ; LOCATION: 1..2889
 ; US-09-442-100-5

RESULTS

US-09-442-100-5

US 03 442 100 3
; Sequence 5, Application US/09442100

QY	3284	ACGCTGGTACGCAAAATCGATACCTCGAACCACTTTGTATGGGATGAAACCACTTCGCGAAA	3343
Db	1654	TGCTTAGCAAGAAAGTCGATA---CTAAGCTTTGTATGCAACAAGAGACTCTTCGAAAG	1710
QY	3344	GCGGACGTTCTCAAGCGGAAATCAAGTGGGCACAGGTGAAGCGCGAGAGGGATATCCTCGCG	3403
Db	1711	AAAGACGTTCTGCTCCGAAATCAAGTGGGCTCATGTGAAAGCGGAGAGGGATATCCTAGCA	1770
QY	3404	GAAGCGGACATAACTGGTGGTCAAGTTCTACTACAGCTTCCAGGCAAGAGATAAATCTG	3463
Db	1771	GAAGCGGAAATGAGTGGGTGGTCCGCTGTACTACTCTTCCAGGCAAGAGCAAACTG	1830
QY	3464	TACTTTGTGATGGACTACATACCAAGTGGTGTCTGATGTGCTGCTCATCAAACTGGCG	3523
Db	1831	TACTTTGTGATGGACTACATTCCTTGGGGGGATATGATGAGGCTATTAAATAGAAATGGC	1890
QY	3524	ATTTTCGAGGAACTGGCCAGATTTCTACATCGCGAGGTCACTGTGGCGCGTGGACAGC	3583
Db	1891	ATCTTTCTCGAAATCTGGCAGATTTCTACATAGCAGAACTTACCTGTGCAAGT	1950
QY	3584	GTTTCAAAAATGGGTTTCATTTCACAGAGACATCAAGCTGACACATCTCATCATGATAGG	3643
Db	1951	GTTTCAAAAATGGGTTTATTATAGAGATATTAAACCTGATAACATTTTGATTCACCGT	2010
QY	3644	GAGCGACATAAAGCTCACCGACTTTGGCTGTGACAGGGATTCGCGATGGACGCAAC	3703
Db	2011	GATGGCCATATTAAATGACTGACTTGGCTTTGTCACCTGGCTTCAGATGGACACATGAC	2070
QY	3704	TCGAAGTACTACAGGAGAACGGCAATCACTCGCGCCAGGACTCGATGGAGCCCTGGGAG	3763
Db	2071	TCCAAGTACTACC---AGAGTGGGATCACCACCGCAAGATAGATGATTCAGTAAC	2127
QY	3764	GAATACTCCGA-----GAAGCGACCGAAGCCCAACCGTGTGGAGAGGCGA	3808
Db	2128	GAATGGGAGATCCTTCCAAATGTCGGTGGGACAGACTGAAGCGACTGGAGCGGAGA	2187
QY	3809	CGGATGGCGGATACCAAGAGTCTGSCCCACTCGCTGGTGGGACCCCGAACTACATA	3868
Db	2188	GCTGCTGCCAGCACCAAGCGATGCTAGGCCCAATCTCTGGTGGGACTCCCAATTAAT	2247
QY	3869	GCTCCGAGGTGCTGGAGAGAGTGGGTACACGAGCTGTGCGACTACTGACGCTGGCG	3928
Db	2248	GCACCTGAAGTGTACTGCGAACAGGATATACACAGCTGTGTGACTGGTGAGCTTGGT	2307
QY	3929	GTACCTCTTATAGATGCTGGTGGGTAGCGCGCCCTTCTTGGCCAAACAGTCCGCTGGAA	3988
Db	2308	GTATTCTTCTGTAATGTTGGTGGGCAACCTCCTTCTTGGCACAACCCCATTAGAA	2367
QY	3989	ACGCAACAAAAGGTCACTCAACTGGGAGAAAACSTGCATATTCGCGCGAGCGCGAGTTA	4048
Db	2368	ACACAAATGAAGTTATCATCTGGCAAACTTCTACACATCCCTCCTCAAGCTAAGCTG	2427
QY	4049	TCGCGGAGGCTACGGACTTGATAAGGAGGCTGTGTGCTGGCTGACAAAGCGCTGGCG	4108
Db	2428	AGTCTCTGAAGGCTCTGACCTCATTTCAAACTGTGTCGAGGACCAAGAACCGCTCGC	2487
QY	4109	AAAG---GCGTGGAGGAGGTCAAGAGCCAGACTTCTTCAAGGGCATCGA---CTTTGGC	4162
Db	2488	AAAGACGGTGTGATGAGATAAAGGCTCATPCCAAATTTTAAAGACCATCGATTTCTAGT	2547
QY	4163	GACATGGCGACCAAAAGCGCCCTACATACCGGAAATCAAGCACCCACGAGACATCC	4222
Db	2548	GATCTGAGACAGCATCTGCTTCATCATCCCTAAAATCAGCATCCCAACAGATACATCC	2607
QY	4223	AACTTTGATCCCGTGGATCCGGAGAGCGTGGCTCGAATGAC	4264
Db	2608	AATTTGACCCCTGTGATCCTGATAAATTTGGAGGCGATGGC	2649

RESULT 6
US-09-442-100-3
; Sequence 3, Application US/09442100
; Patent No. 6359193

GENERAL INFORMATION:

GENERAL INFORMATION:
 APPLICANT: Xu, Tian
 APPLICANT: Tao, Wufan
 APPLICANT: Wang, Weiyl
 APPLICANT: Zhang, Sheng
 APPLICANT: Yu, Wan
 TITLE OF INVENTION: NUCLEOTIDE AND PROTEIN SEQUENCES OF LATS
 NUMBER OF SEQUENCES: 16
 CORRESPONDENCE ADDRESSES:
 ADDRESS: Pennie & Edmonds
 STREET: 1155 Avenue of the Americas
 CITY: New York
 STATE: New York
 COUNTRY: USA
 ZIP: 10036-2711
 COMPUTER READABLE FORM:
 MEDIUM TYPE: Floppy disk
 COMPUTER: IBM PC compatible
 OPERATING SYSTEM: PC-DOS/MS-DOS
 SOFTWARE: Patent In Release #1.0, Version #1.30
 CURRENT APPLICATION DATA:
 APPLICATION NUMBER: US/09/442,100
 FILING DATE:
 CLASSIFICATION:
 PRIOR APPLICATION DATA: 08/411,111
 FILING DATE:
 ATTORNEY/AGENT INFORMATION:
 NAME: Mistock, S. Leslie
 REGISTRATION NUMBER: 18,872
 REFERENCE/DOCKET NUMBER: 6523-003
 TELECOMMUNICATION INFORMATION:
 TELEPHONE: (212) 790-9090
 TELEFAX: (212) 869-9741/8864
 TELEX: 66141 PENNIE
 INFORMATION FOR SEQ ID NO: 3:
 SEQUENCE CHARACTERISTICS:
 LENGTH: 3984 base pairs
 TYPE: nucleic acid
 STRANDEDNESS: double
 TOPOLOGY: unknown
 MOLECULE TYPE: CDNA
 FEATURE:
 NAME/KEY: CDS
 LOCATION: 231..3623
 231-09-442-100-3

Length 3984:

Query Match	8.2%;	Score 470.6;	DB 4;	Length 3984;
Best Local Similarity	62.8%;	pred. No. 1.4e-102;		
Matches 817;	Conservative 4;	Mismatches 453;	Indels 27;	Gaps 5;
QY	2984	AAGGAGAAGGAGGAGGACGCAAGAGTTTCGGCATCAGGCAGTACTCGCGCAAGCCTTTC	3043	
Db	2088	AACAAGAAAGATCAAGAGCGAAGGAAATCTCGTATTCAAAGTTATTCTCTCAAGCATTT	2147	
QY	3044	AAGTCTTCATGGAGCAGCACATAGAGAAGCTGATCAAGTCGTATGCCAGCCACGTAAT	3103	
Db	2148	AAATCTTTATGGAGCAACATGTAGAAAATGTACTCAAATCTCATCAGCAGCTACAT	2207	
QY	3104	CGCAAGAAATCAGCTGGAGAAAGAGATGCACAAAGTTGGGACTGCCCGCATCAGACCCCAAATC	3163	
Db	2208	CGTAAAAACAATTAGAGAAATGAATGATCGGGTTGGATTATCTCAAGATGCCAGGAT	2267	
QY	3164	GAGATGAGGAAAAATGCTGCAACCAAAAGGAGAGCAACTACATTCGATTGAAGCCGCCAAG	3223	
Db	2268	CAATGAGAAGAATGCTTTGCCAAAAAGAAATCTAATTACATCCGCTTTAAAGGGCTTAA	2327	
QY	3224	ATGACAAGAGCATGTTTCGTCAAACATGAAGCCCATTTGGAGTGGGTGCGATTGTGGCAGGTA	3283	
Db	2328	ATGGCAAAAGTCTATGTTTGTGAGATAAAGACACTAGGAATAGGAGCATTTGGTGAAGTC	2387	
QY	3284	ACGCTGGTGGACAAAATPCGATACCTCGAACCAATTTGTATCGCATCAAAACCCCTCGGAAA	3343	


```

Db 2388 TCTAGCAAGAAAGTAGATA---CTAAGGCTTTGTATGCAACAAAACCTCTCGAAAG 2444
Qy 3344 GCGGAGCTTCTCAAGCGGAATCAGGTGGCACACGCTGAAGCGGAGAGGATATCTCGCG 3403
Db 2445 AAGATGTTCTTCTCGGAATCAAGTCGCTCATGTTAAGCGTGAGAGATATCTCGGT 2504
Qy 3404 GAAGCGCAATCAATCAAGTGGTGTGAAGTTGTACTACAGCTTCCAGGACAAGGATAATCTG 3463
Db 2505 GAAGTGCAATGAATGGTAGTCTGCTATATTTATTCATCCAGATAAGGACATTTA 2564
Qy 3464 TACTTTGTAGTACACTACACAGGTGGTGTATGATCGCTGCTCATCAAACTGGGC 3523
Db 2565 TACTTTGTAAATGGACTACATCTCGGGGTGATGATGAGCTTATTAATTAGAATGGC 2624
Qy 3524 ATTTTCGAGGAGAACTGCCAGATCTACATCGCGGAGGTCACTCGCCCTGGACAGC 3583
Db 2625 ATCTTTCGAGAAGTCTGGCAGCTCTACATAGCAGAACTTACCTGTGCACTTGAAGT 2684
Qy 3584 GTTCACAAATGGGCTTCATTCACAGACATCAAGCCTGACAAACATCACTCATGATAG 3643
Db 2685 GTTCATAAATGGGTTTATTCATAGAGATATTAACCTGATAATATTTGATGATCT 2744
Qy 3644 GAGGACACATTAAGCTCACCGACTTGGCCTGTGCGACGGATTCGGATGGACGCACAC 3703
Db 2745 GATGTCATATTAATTAAGTACTGACTTGGCTCTGCACTGGCTTCAGATGGACACAG 2804
Qy 3704 TCGAAGTACTACAGGAGGCAAGCAATCACTCGCCGAGGACTCGATGAGCCCTGGAG 3763
Db 2805 TCTAAGTACTATC---AGAGTGTGACCATCCAGCGCAAGATAGATGATTTAGTAAT 2861
Qy 3764 GAATACTCCGAGAAC-----GGACCGAAGCCCAAGCTGTGCTGAGAGGCGCA 3808
Db 2862 GAATGGGGGATCCCTCAAGCTGTCGATGTGAGACAGACTGAAGCCATTAGCGGAGA 2921
Qy 3809 CGGATGGGGATCACCAAGAGTCTCTGGCCCACTCGCTGGTGGGACCCGCAACTACATA 3868
Db 2922 GCTGACGCGCAGCAGCAGCGATGTAGCACATCTTGGTTGGGACTCCCAATTAAT 2981
Qy 3869 GCTCCCGAGGTCTCGAGAGAGTGGTACACGAGCTGTGCGACTACTGAGAGCGTGGC 3928
Db 2982 GCACCTGAAGTGTGCTACGAAACAGATACACAGTGTGTGATTTGGTGGAGTGTGT 3041
Qy 3929 GTCACTCCYATYAGATGCTGGTGTGAGCGCCCTTCTGCGCAACAGTCCGCTGGAA 3988
Db 3042 GTTATCTTTTGAATGTGTGGGACAACTCTTCTTGGCAACACACCATTAGAA 3101
Qy 3989 AGCAACAAAGGTGATCAACTGGGAGAAACSTGTCATATTCGCGCGAGCGGATTA 4048
Db 3102 ACACAAATGAAGTTTATCAACTGGCAACATCTCTACATTTCCACCAAGCTAAACTC 3161
Qy 4049 TCCCGGAGGTACGAGTGTATAGGAGGTCTGTGCGCTGCGCTGACAGCGGCTGGC 4108
Db 3162 AGCTCTGAAGCTTCGTGATCTTATTAATTAACCTTCCGAGGAGCCGAGATCGCTTAGC 3221
Qy 4109 AAGA---CGTGGAGGAGTCAAGACGACGACTTCTTCAAGGGCATCGACTTTCG 4162
Db 3222 AAGATGCTGTGATGAATAAAGCTCATCATTTTAAACAAATGACTTCTCCAGT 3281
Qy 4163 GACATCGGAGAGAAAGCGCCCTACATPACCGGAAATCAAGCAACCCACGACACATCC 4222
Db 3282 GACCTTGACAGACAGCTGCTTTCATACATTCCTAAATACACACACCAACAGATACATCA 3341
Qy 4223 AACTTTGATCCGCTGGATCCGAGAGGAGTGGCTCGAATGA 4263
Db 3342 AATTTTGATCCCTGTGATCCCTGATAAATTTATGGAGTGTATGA 3382

```

RESULT 7

US-09-328-111-26/c
; Sequence 26, Application US/09328111
; Patent No. 6262333
; GENERAL INFORMATION:

```

; APPLICANT: Endege, Wilson O.
; APPLICANT: Steinmann, Kathleen E.
; APPLICANT: Astle, Jon H.
; APPLICANT: Burgess, Christopher C.
; APPLICANT: Bushnell, Steven E.
; APPLICANT: Carroll III, Eddie
; APPLICANT: Catino, Theodore J.
; APPLICANT: Dertli, Adnan
; APPLICANT: Ford, Donna M.
; APPLICANT: Lewis, Marcia E.
; APPLICANT: Monahan, John E.
; APPLICANT: Schlegel, Robert
; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
; TITLE OF INVENTION: PRODUCTS
; FILE REFERENCE: CCD-257 (US)
; CURRENT APPLICATION NUMBER: US/09/328,111
; CURRENT FILING DATE: 1999-06-08
; EARLIER APPLICATION NUMBER: US 60/088,801
; EARLIER FILING DATE: 1998-06-10
; NUMBER OF SEQ ID NOS: 850
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 26
; LENGTH: 638
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-09-328-111-26

```

Query Match 4.8%; Score 273.6; DB 4; Length 638;

Best Local Similarity 66.4%; Pred. No. 6.4e-56;

Matches 425; Conservative 0; Mismatches 209; Indels 6; Gaps 2;

```

Qy 3074 GTGATCAAGTCTGTATCGCCAGCGCAGCTATCGCAAGAATCAGCTGGAGAGGAGATGCAC 3133
Db 637 GTACTCAATCTCATCAGCAGCGCTCTACATCGTAAAAACAATTAGAGATGAATGATG 578
Qy 3134 AAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGGAAAATGCTGAACCAAAAGGAG 3193
Db 577 CGGGTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTCGCAAAAAGAA 518
Qy 3194 ACCAACTACATTCGATTGAAGCGCCCAAGATGGCAAGAGCATGTTCTCAACTGAAG 3253
Db 517 TCTAATACATCCGCTTTAAAGGGCTAAAATGGACAGTCTATGTTCTGAAGATAAAG 458
Qy 3254 CCATTTGAGTGGGTGCTATTTGGCGAGGTAAACGCTGGTGAAGAAATCGATACCTCGAAC 3313
Db 457 ACATGAGGATGAGGACATTTGCTGAGTCTGTCTAGCAAGAAAGTAGATA---CTAAG 401
Qy 3314 CATTTGTATGCGATGAAACCCCTGCGGAAAGCGGAGCTTCTCAAGCGGAATCAGGTGGCA 3373
Db 400 GCTTTGTATGCAACAAAACCTCTTCGAAATAAAGATGTTCTTCTCGAAATCAAGTCGCT 341
Qy 3374 CAGTGAAGCGCGAGGAGATATCTCGCGGAGCGGACAAATAACTGGGTGGTGAAGTTG 3433
Db 340 CATGTTAAGGCTGAGAGAGATATCTCGGCTGAAGCTGACAATCAATGSGTAGTTCGTCTA 281
Qy 3434 TACTACAGCTTCCAGGACAAGGATATCTGTTGCTGATGAGCTACATACCATGCTGT 3493
Db 280 TATTATTCATCCAGATAGGACAAATTTATCTTTGTAATGACTACATATTCCTGGGGT 221
Qy 3494 GATCTGATGCTGCTCATCAAACTGGGCATTTTTCGAGGAGGAACTGGCCAGATTTAC 3553
Db 220 GATATGATGAGCCTATTAAATTAGAATGGGCATCTTTCCAGAAAGTCTGCGCAGGATCTAC 161
Qy 3554 ATCGCGGAGGTCACTCGCCGCTGGACAGCG---TTCACAAAATGGGCTTCATTACAGA 3610
Db 160 ATAGCAGAACTTACCTGTGCACTTTGAAAGGTGTTCTCTTAAATGGGTTTTTATCATAGA 101
Qy 3611 GACATCAAGCTGACAAACATCTCATGATGAGGAGGACACATAAGCTCACCAGCTTT 3670
Db 100 GATATTAAACCTGATAATATTTGATGATGCTGATGATGATTAATTAATGACTGACTTT 41
Qy 3671 GGCCTGTGCACGGGATTCGGATGGACGCAACACTCGAAGT 3710

```


3348 ACCTTCTCAACGGAATCAGTGGCCACACAGTGAAGCCGAGAGGATATCTCTCGCGAAG 3407
966 ATATGCTTTGAAAACAGACAGGTTGGCCACATTCGTGCGGAGCGTGACATCTAGTGAGG 1025
3408 CCGACAATAAATCGGTGGTGAAGTTGACATACAGCTTCCAGGACAAGGATAATCTGTACT 3467
1026 CAGACAGTTTCTGGTGTGAAAATGTTCTATAGTTTTCAGGATAAGCTAAACCTCTACC 1085
3468 TTGTGATGGATACATACAGGAGTGTGATCTGCTCTCATCAAACTGGGCATTT 3527
1086 TAATCATCGAGTTCTGCTGGAGGGACATGATGACCTTGTGATGAAAACAGACACT 1145
3528 TCGAGGAGAACTGCCAGATTTACATCGCGAGGTCACCTGCGCGGTGACAGCGGTC 3587
1146 TCACAGAGAGGAGACTCAGTTTATATAGCAGAAACAGTATTAGCCATAGACTTATTC 1205
3588 ACAAATGGCTTCAATTCACAGAGACATCAAGCCTGACAACTACTCATCATGATAGGAGC 3647
1206 ACCAATTTGGATTTCATCCAGAGACATCAACACAGACACTCTTTTGGACAGCAAGG 1265
3648 GACACATAAAGCTACCGACTTTGGCCTGTGCACGGGATTCGGATGAGCGACCAACTCGA 3707
1266 GCATGTGAACTTCTGACTTTGGTCTTTGACAGGACTGAAAAACACATAGGACAG 1325
3708 AGTACTACAGGAGAACGGCAATCACTCGCGCCAGGACTCGATGAGCGCTGGGAGGAA 3767
1326 AATTTTA-TAGGAATCTGAACACACAGCCTCCCGCAGTGATTCACCTTCAGAACATGAT 1384
3768 ACTCCGAGAACGAGCGGAGCCACCGTCTCTGGAGAGGCGAGGATGCGCGATCACCAAA 3827
1385 TCACAAAGGAAAGCAGAAACCTTGAAAGAAATAGAGTCA----- 1425
3828 GAGTCTCTGCGCCACTCGCTGGTGGGACCCCGCACTACATAGTCTCCGAGGTGCTGGAGA 3887
1426 ---GCTAGCCTTCTCCACAGTAGGACACTCTGCTGAGGTGATGATGATGATGATG 1481
3888 GRAGTGGTACAGCGAGCTGCGACTACTGGAGCGTGGCGCTCATCTCTTATGAGATGC 3947
1482 AGACCGGGTACAAACAGCTCTGATGTTGGTGGTCTGGGTGATGATGATGATGATG 1541
3948 TGTGGGTACAGCGCGCTTTCTGGCCAAACAGTCCCGTGGAAAGCAACAAAGGTGATCA 4007
1542 TCATCGGCTACCCACCTTTCTGTTCTGAGACCCCTCAAGAGACATATAAGAGGTGATGA 1601
4008 ACTGGGAGAAACSCATATTCGCGCGAGCGCGAGTTATCCGCGAGGCTTACGAGCT 4067
1602 ACTGGAAGAAACCTTTGACATTTCTCCAGAGTTCCTCTCTGAGAAAGCAAGATC 1661
4068 TGATAAGGAGGCT---CTGTGCGTGGCTGACAAAGCGGCTGGCGAGAGCGTGACGAGG 4124
1662 TAATTTGAGGTTCTGCTGTAATGGGAACATAGATTGGAGCTCTCTGAGTTGAGGAAA 1721
4125 TCAAGAGCCAGGACTTCTTCAAGGGCATCGACTTTGCGGACATGCGGAAACAGAAAGCGC 4184
1722 TAAAGGTAACTCTTTTTTGAAGGCGTGGACTGGGAACATATACAGAGAGAGCTGCTG 1781
4185 CCTACATACCGGAATCAAGCAACCCACGAGGACATCAAACTTTGAT 4231
1782 CAATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTCAT 1828

1585 TACTCACCGTCGGTACGACACTACATCGCCCGGAGATCTTCTCTACAGGCTTAC 1644
3899 ACAGAGCTGTGACACTACTGAGGAGTGGCGTCACTCTTAYAGATGCTGTGGGTGAG 3958
1645 GGTACAGAGTGGCGTGGTGGTCCCTCGCGGCGCATGATGACAGTGTCTGATCGGTGG 1704
3959 CCGCCCTTTCTGCCCAACAGTCCGCTGGAAACGCAACAAAGGTCACTAACTGGGAGAAA 4018
1705 CCGCGGTTCTGCTGGAGACCCCGCAGGAACCTACAGGAAGATCATGAATTCGAGCAG 1764
4019 ACSCTGATATTCGGCGGAGGCGGAGTTATCCCGGAGGCTACGAGCTGTGATAGGAGG 4078
1765 AGCTGTGTTCCAGACGACATCCACATCTCATAGGAGCAGAGGACCTCATCCCGCG 1824
4079 CTCTGCTGCTGGTGAACAGCGCTGG---GCAAGAGCGTGGACGAGGTCAAGAGCCAC 4135
1825 CTGCTCTCGGACGCGGAGGAGCGCTCGTGGCATGGCGCCCAACGAATCAAAACCCAC 1884
4136 GACTTCTTCAAGGACATCGACTTTGGCGACATCGGGAAGCAGAAAGCGGCTACATACCG 4195
1885 CCCTTCTCCGCGGTGGACTGGGAGACCATCCGCCAGGTGCGCGCTCCCTACATCCCC 1944
4196 GAAATCAAGCACCTACGAGACATCAACTTTGATCCCGTGGATCGGAGAGGCTGC 4253
1945 AAGCTGTCCAGCGTCAACGACACGCGCTTCTTCCCAACAGCAGCTGGAGACGCTGC 2002

RESULT 11
US-08-860-150-6
; Sequence 6, Application US/08860150B
; Patent No. 5981205
; GENERAL INFORMATION:
; APPLICANT: Hemmings, Brian A.
; APPLICANT: Millward, Thomas A.
; TITLE OF INVENTION: Nuclear DBF2-Related (NDR) Kinases
; FILE REFERENCE: 4-20265/A/PCT
; CURRENT FILING DATE: 1997-06-19
; EARLIER APPLICATION NUMBER: PCT/EP95/05052
; EARLIER FILING DATE: 1995-12-20
; EARLIER APPLICATION NUMBER: 94810746.1
; EARLIER FILING DATE: 1994-12-22
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 3018
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (596)..(1990)
US-08-860-150-6

Query Match 4.1%; Score 232.6; DB 2; Length 3018;
Best Local Similarity 53.1%; Pred. No. 9.6e-46;
Matches 599; Conservative Indels 30; Gaps 4;

3108 AGAATCAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCGATCAGACCCCAATCGAGA 3167
729 AAAAGAAGTTAGAAAAGGTGTAAGAGAGAGGCGCTTAAAGATGAGGAGAAACGACTCC 788
3168 TGAGAAAATGCTGAACCAAAAGAGGACCACTACATTCGATTGAAGCGCGCAAGATGG 3227
789 GGAGATCAGCAGATGCTCGGAGGAAACAGAGTTCTCTGTTTGAAGAGAAACAAAGACTTG 848
3228 ACAAGAGCATGTTGCTCAAACTAGACCCCATTTGAGTGGTGCATTTGGCGAGGTAAACG 3287
849 GATTGGAAGATTTGTAGTCTCTTAAAGTAATAGGACAGAGCATTTGGTGGAGTACGGC 908
3288 TGGTGACAAAATCGATACCTCGAACCATTTGTATGCGATGAAACCCCTCGCGGAAAGCGG 3347
909 TTGTTCAAGACAAAGATA---CGGACATGTTGATGCAATGAAATTAATCTCCGTAAGACG 965

EARLIER FILING DATE: 1997-06-19
EARLIER APPLICATION NUMBER: PCT/EP95/05052
EARLIER FILING DATE: 1995-12-20
EARLIER APPLICATION NUMBER: 94810746.1
EARLIER FILING DATE: 1994-12-22
NUMBER OF SEQ ID NOS: 18
SOFTWARE: Patent in Ver. 2.0
SEQ ID NO 6
LENGTH: 3018
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: CDS
LOCATION: (596)..(1990)
US-09-338-132-6

Query Match
Best Local Similarity 4.18; Score 232.6; DB: 3; Length 3018;
Matches 599; Conservative 2; Mismatches 496; Indels 30; Gaps 4;

Qy 3108 AGAATCAGCTGAGAGGAGATGCAAAAGTGGGACTGCCGATCAGACCCAAATCGAGA 3167
Db 729 AAAAGAACTTGAAGAGTGGAGAGAGAGGCTTAAAGATGAGGAGAACACCTCC 788

Qy 3168 TGAGGAAATGCTGAACCAAGAGAGACAACTACATTCGATTTGAAGCGCGCCAAAGATGG 3227
Db 789 GGAGATCAGACATCTCGGAAGGAAACAGAGTTTCTCGTTTGAAGAGAAACAGACTTG 848

Qy 3228 ACAAGAGATGCTCTCAACTGAAGCCATTTGGAGTGGGTGCTATTTGGCGAGGTAAAGC 3287
Db 849 GATTGGAAGATTTTTCAGTCTTAAAGTAATAGGAGAGGAGCAATTTGGTGAAGTACGCG 908

Qy 3288 TGGTGAGCAAAATCATCTCGAACCACTTTGTATGCGATGAAACCTCGCGAAAGCGG 3347
Db 909 TTGTTCAAGAAAGATA---CGGACATGTATGATCAATGAAATATCTCCGTAAGACAG 965

Qy 3348 ACGTTCTCAAGCGGAATCAGGTGGCGACAGTGAAGCGCGAGAGGATATCTTCGCGGAAG 3407
Db 966 ATATGCTTGAAGAGAGAGGTTGGCCACATTCGTGCGGAGCGTGACATTTAGTGGAGG 1025

Qy 3408 CCGACAATTAAGTGGTGTGAAGTTGTACTACAGCTTCAGGACAGCAAGATATCTGTACT 3467
Db 1026 CAGACAGTTTGTGGGTTGTGAAGATGTTCTATAGTTTTCAGGATAAGTAACCTCTACC 1085

Qy 3468 TTGTGATGAGCTACATACAGGTGGTGTATCTGCTGCTCATCAAACTGGGCAATTT 3527
Db 1086 TAATCATGAGTTCTGCTCGTGGAGGACATGATGACCTTTGTATGAAAAAGACATC 1145

Qy 3528 TCGAGGAGAACTGCCAGATTTACATTCGCGGAGGTACCTGCGGCGCTGGACAGCGTTC 3587
Db 1146 TGACAGAGAGGAGACTCAGTTTATATAGCAAGAAACAGTATTAGCCATAGACTTATTC 1205

Qy 3588 ACAAATGGCTTCAATTCACAGAGACATCAAGCCTGACAACTACTCATCTATAGGGAGC 3647
Db 1206 ACCACTTGGATTCATCCAGAGACATCAACACAGAACCTTTTGGACAGCAAGG 1265

Qy 3648 GACACATAAAGCTCACCAGCTTTGCGCTGTGCAGCGGATTCGATGAGCGACCAACTCGA 3707
Db 1266 GCCATGTGAACCTTCTGACTTTGCTTTTCACAGGACTGAAATAACACATAGGACAG 1325

Qy 3708 AGTACTACAGAGAACGGCAATCACTCGCGCAGGACTCATGATGAGCGCTGGAGGAAT 3767
Db 1326 AATTTTA-TAGGAATCTGAACACAGACGCTCCCGAGTGTACTTTCCAGAACATGAAT 1384

Qy 3768 ACTCGGAGAACGGACCGGACCCACCTGCTGGAGAGCGGATGCGGCATCACCAG 3827
Db 1385 TCCAAAGGAAAGCAGAAACCTGGAAGAAATAGACGTCA----- 1425

Qy 3828 GAGTCTCGCCACTCGCTGTGGGACCCCGAACTACATAGTCCCGAGGTGCTGGAGA 3887
Db 1426 ---GCTAGCCTTCTCCACAGTAGGCACTCTCTAGTACTGCTCTCGTGGAGTTTCATGC 1481

Qy 3888 GRAGTGGGTACACGAGCTGTGCGACTACTGGAGCGTGGCGGTCTATCTTAYGAGATGC 3947

RESULT 13

US-08-878-989-11

Sequence 11, Application US/08878989

Patent No. 5885803

GENERAL INFORMATION:

APPLICANT: Bandman, Olga

APPLICANT: Hillman, Jennifer L.

APPLICANT: Corley, Neil C.

APPLICANT: Guegler, Karl G.

APPLICANT: Lal, Preeti

APPLICANT: Goli, Surya K.

APPLICANT: Shah, Purvi

TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN

NUMBER OF SEQUENCES: 21

CORRESPONDENCE ADDRESS:

ADDRESSEE: Incyte Pharmaceuticals, Inc.

STREET: 3174 Porter Drive

CITY: Palo Alto

STATE: CA

COUNTRY: USA

ZIP: 94304

COMPUTER READABLE FORM:

MEDIUM TYPE: Diskette

COMPUTER: IBM Compatible

OPERATING SYSTEM: DOS

SOFTWARE: FastSeq for Windows Version 2.0

CURRENT APPLICATION DATA:

APPLICATION NUMBER: US/08/878,989

FILING DATE:

CLASSIFICATION: 435

PRIOR APPLICATION DATA:

APPLICATION NUMBER:

FILING DATE:

ATTORNEY/AGENT INFORMATION:

NAME: Billings, Lucy J J

REGISTRATION NUMBER: 36,749

REFERENCE/DOCKET NUMBER: PF-0321 US

TELECOMMUNICATION INFORMATION:

TELEPHONE: 415-855-0555

TELEFAX: 415-845-4166

TELEX:

INFORMATION FOR SEQ ID NO: 11:

SEQUENCE CHARACTERISTICS:

LENGTH: 1935 base pairs

TYPE: nucleic acid

STRANDEDNESS: single

TOPOLOGY: linear

IMMEDIATE SOURCE:

Db 1482 AGACCGGTACAAAGCTCTGTGATTTGGGTGATCATGTATGAGATGC 1541
Qy 3948 TGGTGGGTACGCGGCTTTCTGGCCACAGTCCCTGGAACGCAAAAGGTATCA 4007
Db 1542 TCATCGGTACCCACCTTTCTGTTCTGAGACCCCTCAAGAGACATATAAGAGGTGATGA 1601
Qy 4008 ACTGGGAGAAAACCTGCATATTCGCGCGAGCGGAGTTATCCCGCGAGGCTACGGACT 4067
Db 1602 ACTGGAAGAAACTTTGACTTTTCTCCAGAGTTCCCATCTCTCAGAAAGCAAGGATC 1661
Qy 4068 TGATAAGGAGGCT---CTGTGCGTCTGACAAAGCGCTGGGCAAGACGCTGGACGAGG 4124
Db 1662 TAAATTTGAGGTTCTGCTGTGAATCGGAACATAGAAATGGAGCTCCTCGAGTTGAGGAAA 1721
Qy 4125 TCAAGAGCCAGCACTTCTCAAGGCGCATCGACTTTCCGCGACATGCGGAGACAGAAGCGC 4184
Db 1722 TAAAAAGTAACCTCTTTTGAAGCGCTTGACTGGGAACATATCAGAGAGAGACCTGCTG 1781
Qy 4185 CTACATACCGGAATCAAGCACCCAGCGGACATCATCAACTTTTGTAT 4231
Db 1782 CAATATCTATTGAATCAAGCATTGATGATACCTCAACTTCGAT 1828

LIBRARY: SYNORAT04
CLONE: 705365
US-08-878-989-11

Query Match 3.8%; Score 216.6; DB 2; Length 1935;
Best Local Similarity 50.9%; Pred. No. 5.1e-42;
Matches 609; Conservative 2; Mismatches 556; Indels 30; Gaps 3;

QY 3043 CAAGTCTTCATGAGGAGCAGACATAGAGAACGTCGATCAAGTCGATGCCAGCCAGCTA 3102
DB 121 CAAGCTCACATTTGGAGATTTTATAGCAACCTAATTTTACAGCATGAAGAGAGAGAAC 180
QY 3103 TCSCAAGAAATCAGCTGGAGAGAGAGATGCACAAAGTGGGACTGCCGATCAGACCCCAAT 3162
DB 181 CAGCAGAGAAATTTAGAGTGGCCATGGAAGAAGAGGATTTAGCAGATGAAGAGAAAA 240
QY 3163 CGAGATGAGGAAATGCTGAACCAAGAGGAGAGCAACTACATTCGATTGAAGCCGCCAA 3222
DB 241 GTTACGTCGATCACACACAGCTCCCAAGAAACAGAGTTCTTACGGCTCAAAAGGACAG 300
QY 3223 GATGGCAAGAGCATGTTCCGTCAAACCTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGT 3282
DB 301 ACTTGGCTTGGTACATTTTGGTCTCTCTGAAAGTTATAGGAAGAGAGGCTTTGGAGAGGT 360
QY 3283 AACCTGGTGGAGCAAAATCGATCTCGAACCATTTCTGATCGATGAAGAACCCCTGCGGAA 3342
DB 361 GCGGTGGTCCACAAAAAAGATACAGG---CCATATCTATGCAATGAAGATATTGAGAAA 417
QY 3343 AGCGGAGCTTCTCAAGCGGAAATCAGGTGGCAGACGTCGAAGCCGAGAGGATATCTCCG 3402
DB 418 GTCTGATATGCTTGAAGAGAGAGGTGGCCCATATCCGAGCAGAAAGAGATATTTGGT 477
QY 3403 GGAAGCCGCAATAAATCGGTGGTGAAGTTGTACTACAGCTTCCAGGCAAGGATAATCT 3462
DB 478 AGAAGCAGATGGTCCCTGGGTGGTGAAGATGTTTACAGTTTTCAGGATAAGAGGAATCT 537
QY 3463 GTACTTTGTGATGACTACATACAGGTGGTGTGATCTGATGTGGTCTCATCAAACTGGG 3522
DB 538 TTATCTAATCATGGAATTTCTCCCTGGAGGTGACATGACATTTGTAATGAAGAAAGA 597
QY 3523 CATTTTCGAGGAGAACTGCCAGATTTCTACATCGCGGAGGTCACTCGCCGCTGGACAG 3582
DB 598 CACTTCACAGAGAGAGAAACAGATTTCTACATTTACAGACTGTTCTGCAATAGATGC 657
QY 3583 CGTTCACAAATGGGTTCATTCACAGAGACATCAAGCCCTGACAACATCTCATCGATAG 3642
DB 658 GATCCACCACTGGGTTCATCTCATCGGATATTAAAGCCAGACAACTTTTATTGGATGC 717
QY 3643 GGAGGACACATAAGCTCACCGACTTTGGCCGTGTGCACCGGATTCGATGGAGGCACAA 3702
DB 718 CAAGGGTCAATGTAATATCTGATTTTGGTTCATGTACGGGATTAAGAAAGCTCACAG 777
QY 3703 CTCGAAGTACTACAGAGAGAGCAATCACTCGCGCAGACTCGATGGAGCCCTGGGA 3762
DB 778 GACTGAATTTTATAGAAATCTCACACACACCCACCAAGTACTTCTCATTTTCAGAACAT 837
QY 3763 GGAATACTCCGAGAGAGGACCGGAGCCACCGTGTGTGGAGAGCGGACGGATCGCATCA 3822
DB 838 GAACTCAAGAGAGGAGAGAACTTGGAAAGAGAAAGAGAGACAA-----883
QY 3823 CCAAGAGTCTCGCCCACTCGCTGGTGGCCACCCGAACTACATAGTCCCGAGGTGCT 3882
DB 884 -----CTGGCATTTTCCACAGTTGGGACACCAAGATTACATTTGCTCCAGAAATTT 933
QY 3883 GGAGAGAGTGGGTACACGAGCTGTGCGACTACTGGAGCGTGGCGGTCTATCTCTTAYGA 3942
DB 934 CATGACAGCTGGTTACACAAATTTGTGACTGTGTGTTCTTGGAGTGTATGTATGA 993
QY 3943 GATGCTGGTGGGTGACGCGCCCTTTCTGGCCAAACAGTCCGCTGGAAACCGCAACAAAGGT 4002
DB 994 AATGCTAATAGGATATCCACTTTCTGCTGTGAACACACCTCAAGAAACATACAGAAAAAGT 1053
QY 4003 CATCAACTGGGAGAAACCTGTGATATTCGCGCGAGCGGAGTATTCGCGGAGGCTAC 4062

DB 1054 GATGAAGTGAAGAGAACTCTGTGTTATTTCTCCAGAGGTACCTATATCTGAGAAAGCCAA 1113
QY 4063 GGACTTGTATAGGAGGCTCTGTGCTGGCTGCAAGCGCTGGGCAAGA---CGGTGGA 4119
DB 1114 GGACTTAAATCTCAGATTTTGTATTTGATTCGAAAACAGAAATTTGGAATATGAGGTAGA 1173
QY 4120 CGAGGTCAAGAGGACCACTTTCTTCAAGGGCATCGACTTTTGGGACATTCGCGAAGACGAA 4179
DB 1174 AGAAATAAAGGTCTATCCCTTTTGAAGGTGTCGACTGGGAGCACATAAGGGAAGGCC 1233
QY 4180 AGCCCTTACATACCGGAATCAAGCAACCCACGACACATCCAACTTTGATCCCGT 4236
DB 1234 AGCAGCAATCCCTATAGAAATCAAAAGCAATGATGACTTTCAAAATTTGATGACTT 1290

RESULT 14

US-09-272-796-11
; Sequence 11, Application US/09272796
; Patent No. 6207148
; GENERAL INFORMATION:
; APPLICANT: Bandman, Olga
; APPLICANT: Hillman, Jennifer L.
; APPLICANT: Corley, Neil C.
; APPLICANT: Guegler, Karl G.
; APPLICANT: Lal, Preeti
; APPLICANT: Goli, Surya K.
; APPLICANT: Shah, Purvi
; TITLE OF INVENTION: DISEASE ASSOCIATED PROTEIN
; NUMBER OF SEQUENCES: 21
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Incyte Pharmaceuticals, Inc.
; STREET: 3174 Porter Drive
; CITY: Palo Alto
; STATE: CA
; COUNTRY: USA
; ZIP: 94304
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; OPERATING SYSTEM: DOS
; SOFTWARE: FASTSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/272,796
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 08/878,989
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Billings, Lucy J J
; REGISTRATION NUMBER: 36,749
; REFERENCE/DOCKET NUMBER: PF-0321 US
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 415-855-0555
; TELEFAX: 415-845-4166
; TELEX:
; INFORMATION FOR SEQ ID NO: 11:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1935 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; IMMEDIATE SOURCE:
; LIBRARY: SYNORAT04
; CLONE: 705365
; US-09-272-796-11

Query Match 3.8%; Score 216.6; DB 4; Length 1935;
Best Local Similarity 50.9%; Pred. No. 5.1e-42;
Matches 609; Conservative 2; Mismatches 556; Indels 30; Gaps 3;

QY 3043 CAAGTCTTTCATGGAGCAGCATAGAAACCTGATCAAGTCTGATCGCCAGCCACCTA 3102
 DB 121 CAAGCTCATATGGAGATTTTATAGCAACCTAATTTTACAGCATGAAGAGAGAGAA 180
 QY 3103 TCGAAGATACGTGGAGAGAGATGCAAAAGTGGAGTCCGATGACACCAAT 3162
 DB 181 CAGGAGAGAAATAGAGTGGCATGGAAGAAGGATTTAGCAGATGAAGAGAA 240
 QY 3163 CGAGATGAGAAATGCTGAACCAAGAGAGCAACTACATTCGATTGAAGCGCCAA 3222
 DB 241 GTTACGTGATCACAACAGCTCGCAAGAAACAGATTCACGGCTCAAAAGACCA 300
 QY 3223 GATGACAAGAGCATGTTGCTGAACCTGAAGCCCTTGGAGTGGTGGTGGAGGT 3282
 DB 301 ACTTGGCTGGATGACTTGGTCTCTGAAGTCTATAGGAAGAGGAGTCTTTGGAGAGT 360
 QY 3283 RACGCTGGTGAACAAATCGATACCTCGAACCATTTGATGGAGTGAACACCTCGGAA 3342
 DB 361 GCGGTGGTCCCAAAAGATACAGG--CCATATCTATGCAATGAAGATATTGAGAA 417
 QY 3343 AGCGAGCTTCTCAAGCGGAATCAGGTGGCAGACCTGAAGCGGAGAGGATATCTCGC 3402
 DB 418 GTCGATATGTTGAAGAAGCAGGTGGCCATATCGAGCAGAAAGAGATATTTGGT 477
 QY 3403 GGAAGCCGACATAACTGGTGGTGAAGTTGTACTACAGCTTCCAGGACAGGATAATCT 3462
 DB 478 AGAAGCAGATGGTGGTGGTGAAGATGTTTACAGTTTTCAGGATAGAGGAATCT 537
 QY 3463 GTACTTTGTGATGATACATACAGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 3522
 DB 538 TTATCTAATCATGGAATTTCTCCCTGGAGGTGACATGATGACATTTGCTAATGAAGAA 597
 QY 3523 CATTTTGGAGGAGAACTGGCCAGATCTTACATCGCCGAGGTCACTCGCCGCGGACAG 3582
 DB 598 CACCTTGACAGAGAGAAACACAGTTCTACATTTACAGAGACTGTCTGGCAATGATGC 657
 QY 3583 CTTTCAAAAATGGCTTTCATTCACAGACATCAAGCTGACACATCTCATCTGATAG 3642
 DB 658 GATCCACAGTGGTGGTTCATCCATCGGATATTAAGCCAGACACCTTTTATGGATGC 717
 QY 3643 GGAGGACACATAAGCTCACCGACTTTGGCTGGTGGTGGTGGTGGTGGTGGTGGTGG 3702
 DB 718 CAAGGTGATGATAAATATTCTGATTTTGGTTCATGTGACGGGATTAAGAAAGCTCACAG 777
 QY 3703 CTCGAAGTACTACAGGAGAACGGCAATCACTCGCCGAGGACTCGATGGAGCCCTGG 3762
 DB 778 GACTGAATTTATAGAAATCTCACACACACCCCAAGTACTCTCTATTTCAAGACAT 837
 QY 3763 GGAATACCTCGAGAACGGAAGCCACCGTCTGGAGAGCGGACGATGCGCGATCA 3822
 DB 838 GAACCTAAGAGAGAAAGCAGAACTTGAAGAAGAAACAGGAGACAA----- 883
 QY 3823 CCAAGAGTCTGGCCACTCGTGGTGGGACCCGCACTACATAGTCTCCGAGGTGCT 3882
 DB 884 -----CTGGGATATTCCACAGTTGGGACACCATGATTTGCTCCAGAAAGTATT 933
 QY 3883 GGAGAGAGTGGGTACAGCAGCTGTGCGACTACTTGGAGCGGTGGCGCTCATCTTATGA 3942
 DB 934 CATGAGACTGGTTACACAATTTGTGACTGGTGGTGGTGGTGGTGGTGGTGGTGGTGG 993
 QY 3943 CATGCTGGTGGGTACGCGCCCTTTCTGGCCCAACAGTCCGCTGGAAACGCAACAAAGGT 4002
 DB 994 AATGCTAATAGATATCCACCTTTCTGCTCTGAACACCTCAAGAAACATACAGAAAGT 1053
 QY 4003 CATCAACTGGGAGAAACCTGTCATATTTCCGGCCGAGCGGAGTTATCCGGGAGGCTAC 4062
 DB 1054 GATGAACCTGGAAGAAACCTGTTGTTTCTCCAGAGGTACCTATATCTGAGAAAGCCAA 1113
 QY 4063 GGACTTGAAGAGGCTCTGTGCGCTGACAAAGCGGTGGGCAAGA--GGTGGGA 4119
 DB 1114 GGACTTAATCTCAGATTTTGTATTGATCTGAAACAGAAATGGAATAGTGGAGTAGA 1173
 QY 4120 CGAGTCAAGAGCCACGACTTCTTCAAGGGGATCTGACTTTTGGGACATCGCGAAGCAGAA 4179

DB 1174 AGAAATAAAGTCTATCCCTTTTGAAGTGTGACTGGGAGCACAATAAGGAAAGGCC 1233
 QY 4180 AGCGCCCTTACATACCGGAAATCAAGCACCCRAGGACACATCCAACTTTGATCCCGT 4236
 DB 1234 AGCAATATCCCTATAGAAATCAAAAGCATTTGATGATATCTTCAAAATTTGATGACTT 1290

RESULT 15
 US-09-328-111-66
 ; Sequence 66, Application US/09328111
 ; Patent No. 6262333
 ; GENERAL INFORMATION:
 ; APPLICANT: Endege, Wilson O.
 ; APPLICANT: Steinmann, Kathleen E.
 ; APPLICANT: Astie, Jon H.
 ; APPLICANT: Burgess, Christopher C.
 ; APPLICANT: Bushnell, Steven E.
 ; APPLICANT: Carroll III, Eddie
 ; APPLICANT: Catino, Theodore J.
 ; APPLICANT: Derti, Adnan
 ; APPLICANT: Ford, Donna M.
 ; APPLICANT: Lewis, Marcia E.
 ; APPLICANT: Monahan, John E.
 ; APPLICANT: Schlegel, Robert
 ; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
 ; FILE REFERENCE: CCD-257 (US)
 ; CURRENT APPLICATION NUMBER: US/09/328,111
 ; CURRENT FILING DATE: 1999-06-08
 ; EARLIER APPLICATION NUMBER: US 60/088,801
 ; EARLIER FILING DATE: 1998-06-10
 ; NUMBER OF SEQ ID NOS: 850
 ; SOFTWARE: FastSeq for Windows Version 3.0
 ; SEQ ID NO 66
 ; LENGTH: 678
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; FEATURE:
 ; NAME/KEY: misc_feature
 ; LOCATION: (1)...(678)
 ; OTHER INFORMATION: n = A,T,C or G
 US-09-328-111-66

Query Match 2.9%; Score 164.6; DB 4; Length 678;
 Best Local Similarity 63.2%; Pred. No. 7.7e-30;
 Matches 289; Conservative 0; Mismatches 161; Indels 5; Gaps 2;

QY 3078 TCAAGTCTGATCGCCAGCGCAGCTATCGCAAGAATCAGCTGGAGAGGAGATGCAACAAG 3137
 DB 3 TCAATCTCATCAGCAGCGCTCTACATCGTAAACAAATAGAGAATGAATGTCGGG 62
 QY 3138 TGGGACTCCCGATCAGACCCCAATCGAGTGGAGAGAAATGCTGAACCAAAAGGAGAGA 3197
 DB 63 TTGGATTATCTCAAGATGCCAGGATCAATGAGAAAGATGCTTTGCCAAAAGAACTTA 122
 QY 3198 ACTACATTCGATTGAAGCGCGCAAGATGGACAAGAGATGTTCTGCTCAAACTGAAGCCCA 3257
 DB 123 ATTACATCCGCTTAAAAGGCTTAAATGGCAAGTCTATGTTTGTGAAGATAAAGACAC 182
 QY 3258 TTGGAGTGGGTGATTTTGGCGAGGTAAAGCTGGTGGAGCAAAATCGATACCTCGAACCAT 3317
 DB 183 TAGGAATAGGAGCATTTTGGTGAAGTCTGCTAGCAAGAAAGTAGATA--CTAAGGCTT 239
 QY 3318 TGTATGCGATGAACCCCTCGGGAAGCGGAGTCTTCAAGCGGAATCAGGTGGCACAG 3377
 DB 240 TGTATGCAACAAAACCTCTTGGAAAGAAAGATGTTCTTCTTGGAAATCAAGTCGCTCATG 299
 QY 3378 TCAAGCGCGAGGAGATATCTCGCGGAAGCGCAATAAATCTGGTGGTGGTGGTGGTGGT 3437
 DB 300 TTAAGGCTGAGAGAGATATCTTGGCTGAAGCTGACATGAATGAATGGTAGTCTGCTATATT 359
 QY 3438 ACAGCTTCCAGGACAAGGATAATCTGTACTTTGTGTGAGACTATACATACCAGGTGGTGATC 3497

Fri Jan 17 11:17:10 2003

us-09-763-334-7.rni

Page 16

Db 360 ATTCATTCCAAGATAAGGCCATTATCCTT-GTAATGGCTACATTCCTNGGGTGATA 417
QY 3498 TGATGTCGCTGCTCATCAAACTGGGCATTTT 3528
Db 418 TGAAGAGCCCAATTAATANAATGGGCATCTT 448

Search completed: January 16, 2003, 21:59:35
Job time : 245.379 secs

GenCore version 5.1.3.
Copyright (c) 1993 - 2003 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2003, 09:56:42 ; Search time 158.731 Seconds
(without alignments)
16067.254 Million cell updates/sec

Title: US-09-763-334-7
Perfect score: 5720
Sequence: 1' atctgacgacgagcagca.....gccttgccaattagtattac 5720

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 393868 seqs, 222934149 residues
Total number of hits satisfying chosen parameters: 787736

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications_NA:
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq.*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq.*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq.*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq.*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq.*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq.*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq.*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq.*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq.*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq.*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq.*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq.*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq.*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Query Match	Score	Length	DB ID	Description
1	487.8	8.5	2043	9	US-09-836-392-2
2	321.4	5.6	676	9	US-09-764-868-214
3	273.6	4.8	638	10	US-09-879-536-26
4	232.6	4.1	3583	9	US-09-974-238-152
5	200.6	3.5	1452	9	US-09-938-842A-2402
6	195.8	3.4	1689	9	US-09-938-842A-1861
7	188.6	3.3	1818	10	US-09-771-161A-89
8	164.6	2.9	678	10	US-09-879-536-66
9	161.2	2.8	734	9	US-09-764-868-196
10	144.6	2.5	1398	9	US-09-938-842A-633
11	138.6	2.4	1416	9	US-09-938-842A-2503
12	134.6	2.4	3061	10	US-09-880-107-2146
13	134.6	2.4	3407	10	US-09-971-845-1
14	133	2.3	362	9	US-09-796-692-7789
15	127	2.2	1635	10	US-09-880-107-2340
16	121.8	2.1	1244	10	US-09-771-161A-38
17	121.8	2.1	1393	10	US-09-771-161A-37
18	121	2.1	1735	9	US-09-764-868-58
19	120.6	2.1	2549	10	US-09-880-107-3691

20	117	2.0	2556	10	US-09-817-310-1	Sequence 1, Appli
21	116.6	2.0	1257	10	US-09-799-875-15	Sequence 15, Appl
22	116.6	2.0	1826	10	US-09-799-875-13	Sequence 13, Appl
23	116.6	2.0	1972	9	US-10-098-841-166	Sequence 166, App
24	115.4	2.0	2146	10	US-09-954-456-527	Sequence 527, App
25	115.4	2.0	2146	10	US-09-842-307-1	Sequence 1, Appli
26	114.4	2.0	1479	10	US-09-771-161A-46	Sequence 46, Appl
27	113.4	2.0	2637	10	US-09-799-875-3	Sequence 3, Appli
28	113.4	2.0	3003	10	US-09-799-875-1	Sequence 1, Appli
29	113	2.0	277	10	US-09-294-093B-2043	Sequence 2043, Ap
30	110.8	1.9	568	9	US-09-796-692-8122	Sequence 8122, Ap
31	110.8	1.9	6409	10	US-09-864-864-293	Sequence 293, App
32	110	1.9	2610	10	US-09-970-000-3	Sequence 3, Appli
33	104.8	1.8	1812	10	US-09-971-118-1	Sequence 1, Appli
34	104.6	1.8	1461	10	US-09-771-161A-63	Sequence 63, Appl
35	104.6	1.8	1515	10	US-09-804-471A-1	Sequence 1, Appli
36	104.6	1.8	5877	12	US-10-028-946-3	Sequence 3, Appli
37	104.6	1.8	6165	12	US-10-028-946-1	Sequence 1, Appli
38	104	1.8	3244	9	US-10-174-590-571	Sequence 571, App
39	104	1.8	3244	9	US-10-176-758-571	Sequence 571, App
40	104	1.8	3244	12	US-10-052-586-571	Sequence 571, App
41	103.8	1.8	6303	9	US-09-974-238-168	Sequence 168, App
42	102.8	1.8	512	10	US-09-867-701-6230	Sequence 6230, Ap
43	102.2	1.8	2519	9	US-10-071-766-13	Sequence 13, Appl
44	102.2	1.8	2557	10	US-09-972-694-1	Sequence 1, Appli
45	101.4	1.8	678	10	US-09-841-683-4	Sequence 4, Appli

ALIGNMENTS

RESULT 1
US-09-836-392-2
; Sequence 2, Application US/09836392
; Patent No. US20020173458A1
; GENERAL INFORMATION:
; APPLICANT: Ruben et al.
; TITLE OF INVENTION: Protein Tyrosine Kinase Receptor Polynucleotides, Polypeptides
; FILE REFERENCE: PT020P1
; CURRENT APPLICATION NUMBER: US/09/836,392
; CURRENT FILING DATE: 2001-04-18
; PRIOR APPLICATION NUMBER: PCT/US00/28066
; PRIOR FILING DATE: 2000-10-11
; PRIOR APPLICATION NUMBER: 60/159,542
; PRIOR FILING DATE: 1999-10-15
; PRIOR APPLICATION NUMBER: 60/165,914
; PRIOR FILING DATE: 1999-11-17
; PRIOR APPLICATION NUMBER: 60/189,027
; PRIOR FILING DATE: 2000-03-14
; NUMBER OF SEQ ID NOS: 34
; SOFTWARE: PatentIn ver. 2.0
; SEQ ID NO 2
; LENGTH: 2043
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-836-392-2

Query Match 8.5%; Score 487.8; DB 9; Length 2043;
Best Local Similarity 65.2%; Pred. No. 1.3e-97;
Matches 819; Conservative 4; Mismatches 401; Indels 33; Gaps 6;

QY	3130	GCACAAAGTGGGACATCCGCGATCAGACACCAATCGAGATGAGAAATGCTGCTCAAACT	3189
DB	211	CGGTAATGCTGCGATCTCTGAAGCTGAGCAGGAGGATCGGGAAGATCCTCTACAGAA	270
QY	3190	GGAGAGCACTACATTCGATTGAAGCGCCCAAGATGACACAGCATGCTTCTGCTCAAACT	3249
DB	271	AGAGTCTTATACACAGGTTAAAGAGGCCCAAGATGACACAGTCTATGTTTGTCAAGT	330
QY	3250	GAAGCCCATGAGGTGGTGCATTTGGCGAGGTAAAGCTGGTGGAGCAAAATCGATACCTC	3309
DB	331	CAAAACCCCTGGGATCGGTCCTTTGGAGAACTGTGCTTGTGAAGGTGA---CAC	387

Fri Jan 17 11:17:11 2003

```

3310 GAACCAATTTGATCGATGAAACCTCGGAAAGCGGACGTTCTCAAGCGAATCAGT 3369
3380 TCACGCCCTGTACGCCATGAAGACCTTAAGAAAAGGATGCTCTGAACCGGAATCAGT 447
3370 GGCACAGCTGAAGCGCGAGAGGATATCTCTCGCGGAGCGACATAAATCTGGTGGTGA 3429
448 GGCACAGCTGAAGCGCGAGAGGATATCTCTCGCGGAGCGACATAAATCTGGTGGTGA 507
3430 GTTGTACTACAGCTTCAGACAGAGGATATCTCTCTTGTGATGAGTACATACACAG 3489
508 ACTCTACTACTCTCTCAAGCAAGAGGACGCTGTACTTTGTGATGAGTACATCTCTGG 567
3490 TGGTGTATCTGATGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGAACTGGCCAGATT 3549
568 TGGGACATGATGAGCTGCTGATCCGATGGAGTCTTCCCTGAGCAGCTGGCCGGTT 627
3550 CTACATCCCGAGGTACCTGCGCGGTGGACAGGTTTCAAAAATGGCTTCAATCACAG 3609
628 CTACATCCCGAGGTACCTGCGCGGTGGAGTGTCCACAGATGGCTTCAATCACAG 687
3610 AGACATCAAGCTGACATCACTCATCGATAGGAGCGGACACATAAAGCTCACCGACT 3669
698 AGACATCAAGCTGATACATTTTGTAGATCTGTGATGCTCAATTAATCAACATCA 747
3670 TGGCCTGTGCACGGGATTCGATGAGCGACCAACTCGAAGTACTACAGGATTTTACC 3723
748 CGGCTCTGCACTGGGTTTCAAGTGGACTTCAATTTCCAAATATTACCAAGAGGAGCC 807
3724 CGGCAATCACTCGCGCAGGAGCTGATGAGCGCTGGGAGGAACTACTCCGA-----GAA 3777
808 TGTACAGACAGGACGATGGAGCCAGGACCTCTGGGATGATGTGCTAAGTGTGGTG 867
3778 CGGACCAAGCCACCGCTGCTGGAGGCGAGCGATGCGGATACCCAAAGTCTCTGGC 3837
868 TGGGACAGGCTGAAGACCTTAGACAGAGGCGGAGCAGCAGGAGGCTCTCCCGAAGGTA 987
3838 CCACTCGCTGGTGGGACCGGCTGATGAGTGTGCTGAGTGTGCTGAGAGTGGGTA 3897
928 ACATTCACCTGGTGGGACCTGCAATCTATCATCCACCGGAGTGTCTCCCGAAGGTA 987
3898 CAGCAGCTGTGGGACTACTGGAGCGTGGGCTGATCTCTTATGAGATGCTGGTGGTCA 3957
988 CACTCACTGTGACTGGTGGAGTGTGGAGTGTGCTGAGATGCTGCTGGGAGCA 1047
3958 GCGCGCTTTCTGGCCAAAGCTCGGCTGGAAAGCGCAAAAGGCTCATCACTGGGAGAA 4017
1048 GCGCGCTTTTGGCACCCTACTCCCAAGAACCCAGCTGAAGTGAAGTCACTGGGAGAA 1107
4018 AAGCTGATATTCGCGCGAGCGGAGTTATCCCGGAGGCTACGGAGCTTGAAGGAG 4077
1108 CAGCTCCCATTTCCAGCCCGAGTGAAGCTGAGCCCTGAGCGGAGGAGCTTCAACCAA 1167
4078 GCTGTGCGTGGCTGACAGCGGCTGGGAGAGCGGAGGAGGCTTCAAGAGGCA 4134
1168 GCTGTGCTGCTCGCAGACAGCGCTGGGCGGAGTGGGCGGAGTGAAGCTTGAAGCCCA 1227
4135 CGACTTCTCAAGGCACTGACTTTGCG---GGACATCGCGAAGCAAGAAAGCGGCTTACAT 4191
1228 CCCCTTCTTCAAGCGGATGACTTCTCCAGTGACATCCGGAAGCAGCAGCGGAGGAGCT 1287
4192 ACCGGAATCAAGCCCGAGGACATCAACTTTGATCCGCTGATCCGAGAGGAGCT 4251
1288 TCCCAACATCAGCAGCCCGAGGAGCTGGAATTTGAGCCCGGATGAGTGAAGAGCC 1347
4252 GCGCTCGAATCACTCCACATGAGCAGCGGAGTGTGTC-----GACCAGAA 4299
1348 TTGGAAAGCTGACAGGAGTGAAGCAGGCTTGGACACACTACCTCGGCCAATAA 1407
4300 TGACCGCAGCTTTCACGCGCTTTTTCGAATTTACCTTCGCTGGCTTCTTCGACGCA 4356
1408 CAAGCATCTGAGCAGGCTTTTACGAATTCACCTTCCGAGGTTCTTTGATGACAA 1464

```

```

RESULT 2
US-09-764-868-214
; Sequence 214, Application US/09764868
; Patent NO. US20020168711AI
; GENERAL INFORMATION:
; APPLICANT: Rosen et al.
; TITLE OF INVENTION: Nucleic Acids, Proteins, and Antibodies
; FILE REFERENCE: PT232
; CURRENT APPLICATION NUMBER: US/09/764,868
; CURRENT FILING DATE: 2001-01-17
; Prior application data removed - refer to PALM or file wrapper
; NUMBER OF SEQ ID NOS: 1510
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 214
; LENGTH: 676
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: SITE
; LOCATION: (628)
; OTHER INFORMATION: n equals a,t,g, or c
US-09-764-868-214

Query Match 5.6%; Score 321.4; DB 9; Length 676;
Best Local Similarity 69.7%; Pred. No. 2.8e-61;
Matches 474; Conservative 2; Mismatches 199; Indels 5; Gaps 3;

QY 3041 TTCAAGTCTTCTCATGAGCAGCACATAGAGAACGTGATCAAGTCTGATCGCCAGCGCAG 3100
Db 1 TTTAAGTCTTCTCATGAGCAGCACATAGAGAACGTGATCAAGTCTGATCGCCAGCGCAG 60
QY 3101 TATCGCAAGATCAGCTGGAGAGGAGATGACAAAGTGGGACTCCCGATCAGACCCAA 3160
Db 61 AACCGGAGGCTGAGCTGGAGCAAGAAATGGCCAAAGCTGGACTCTGTGAAGCTGAGCAG 120
QY 3161 ATCGAGATGAGAAAATGCTGACCAAAAAGAGGAGCAACTATCTGATGAGTGAAGCGGCC 3220
Db 121 GACCAAGTGGGAAGTCTCTTACCAGAAAGAGTCTTAATACACAGGTAAAGAGGAGGCC 180
QY 3221 AAGATGAGCAAGAGCTGCTGCAAACTCAAGCCCATTTGGAGTGGTGCATTTGGCGAG 3280
Db 181 AAGATGAGCAAGTCTGTTTGTCAAGATCAAAACCTGGGATCGTGCCTTTGAGAGAA 240
QY 3281 GTACGCTGTGAGCAAAATCGATACCTCGAACCATTTGTATGCGATGAACCCCTGCGG 3340
Db 241 GTGTCCTTCTGTTGAAGTGA---CACTACGCGCTGTACGCCATGAAGACCCCTAAGG 297
QY 3341 AAACCGGAGCTTCTCAAGCGGATCAGGTGGCACAGTGAAGCCGAGAGGAGGATATCCTC 3400
Db 298 AAAAAGGATGCTCTGAACCGGAATCAGGTGGCCCAAGCTCAAGCCGAGAGGAGGATCCTG 357
QY 3401 CGGGAAGCGCAATAAATCTGGTGGTGAAGTGTGTACTTACAGCTTCCAGGACAAGGATAT 3460
Db 358 GCCGAGGAGCAAAATGAGTGGTGGCAACTC-TACTACTCTTCCAGAGCAAAAGAGCAG 416
QY 3461 CTGTACTTTGTGTGAGTACATACAGAGTGGTGTATCTGATGCTGCTGCTCAACAACTG 3520
Db 417 CTGTACTTTGTGTGAGTACATCCCTGGTGGGAGATGATGAGCCTTGTGATCCGAGTG 476
QY 3521 GGCATTTTCGAGGAGAACTGGCAGATTTCTACATCGCCGAGGTACCTCGCGCGTGGAC 3580
Db 477 GAGGTCTTCCCTGAGCAGCTGGCCGKCTACATCGCARA-CTGACTTTGGCCATTTGAG 535
QY 3581 AGCGTTCAAAAATGGGTTCAATTCAGAGACATCAAGCCTGACACATACATCATCATGAT 3640
Db 536 AGTGTCCACAAGATGGGTTTATCCACCGAGACATCAAGCCTGATTAACATTTGATGAT 595
QY 3641 AGGAGCGGACACATAAAGCTCACCGACTTTGGCTGTGACCGGATTTCCGATGAGCGCAC 3700
Db 596 CTGGATGTGTACATTAACATCACAGATTTCCGNCCTCTGCAGTGGTTCAGAGTGGACTCAC 655
QY 3701 AACTCGAAGTACTTACCAGGA 3720

```

Db 656 ATATTCACATATTACAGAA 675

RESULT 3

US-09-879-536-26/c

; Sequence 26, Application US/09879536

; Patent No. US20020144298A1

; GENERAL INFORMATION:

; APPLICANT: Endege, Wilson O.

; APPLICANT: Steinmann, Kathleen E.

; APPLICANT: Astle, Jon H.

; APPLICANT: Burgess, Christopher C.

; APPLICANT: Bushnell, Steven E.

; APPLICANT: Carroll III, Eddie

; APPLICANT: Catino, Theodore J.

; APPLICANT: Derti, Adrian

; APPLICANT: Ford, Donna M.

; APPLICANT: Lewis, Marcia E.

; APPLICANT: Monahan, John E.

; APPLICANT: Schlegel, Robert

; TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION

; TITLE OF INVENTION: PRODUCTS

; FILE REFERENCE: CCD-257 (US)

; CURRENT APPLICATION NUMBER: US/09/879,536

; CURRENT FILING DATE: 2001-09-21

; PRIOR APPLICATION NUMBER: US 60/088,801

; PRIOR FILING DATE: 1998-06-10

; NUMBER OF SEQ ID NOS: 850

; SOFTWARE: FastSeq for Windows Version 3.0

; SEQ ID NO 26

; LENGTH: 638

; TYPE: DNA

; ORGANISM: Homo sapiens

US-09-879-536-26

Query Match 4.8%; Score 273.6; DB 10; Length 638;
Best Local Similarity 66.4%; Pred. No. 8.8e-51;
Matches 425; Conservative 0; Mismatches 209; Indels 6; Gaps 2;

QY 3074 GTGATCAAGTCGTATCGCCGCGGACGATATCCAGAGATCAGCTGGAGAGAGATGCAC 3133

Db 637 GTATCAATCTCATCAGCAGCGTCTACATCGTAAACAAATAGAGAAATGAAATGATG 578

QY 3134 AAATGGGATCCCGATCAGACCAATCGAGATGAGGAAATGCTGAACCAAAAGGAG 3193

Db 577 CGGTGGATATCTCAAGATGCCAGGATCAATGAGAGAGATGCTTGGCCAAAAGAA 518

QY 3194 AGCAACTACATTCGATGAAGCGGCCCAAGATGGACAGAGATGTCGTCAAACTGAAG 3253

Db 517 TCTAATTACATCCGCTCTAAAGGGCTAAATGGACAAAGTCTATGTTGTGAAGATAAG 458

QY 3254 CCCATTGGAGGGTGCATTTGGCGAGGTAAAGCTGGTGAAGAAATCGATACCTCGAC 3313

Db 457 ACATGAGGAATAGGAGCATTTGGTGAAGTCTCTAGCAAGAAAGTAGATA---CTAAG 401

QY 3314 CATTTGTATGCGATGAAACCCCTGCGGAAAGCGGACGTTCTCAAGCGGAATCAGGTGGCA 3373

Db 400 GCTTTGTATGCAACAAACTCTCGAATAAGATGTTCTTCGAAATCAAGTCGT 341

QY 3374 CACGTGAAGCGGAGAGGATATCTCGCGGAGCGGACAAATCACTGGTGGTGAAGTTG 3433

Db 340 CATGTTAAGGCTGAGAGAGATATCTGGCTGAAGCTGACAAATGAATGGTAGTCTCTA 281

QY 3434 TACTACACTCCAGGACAGGATATCTGTACTTTGTAGTACGATACATACAGGTGGT 3493

Db 280 TATTATTCATCAAGATAAGGACAAATTTATCTTTGTAAGGACTACATTCCTGGGGT 221

QY 3494 GATCTGATGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGAACTGGCCAGATTCTAC 3553

Db 220 GATATGATGAGCCTATTATTAATGATGGGATCTTCCAGAAAGTCTGCCAGCATCTAC 161

QY 3554 ATCGCCGAGGTACCTGCGCGCTGGACAGCG---TTCACAAATGGGCTTCATTCACAGA 3610

Db 160 ATAGCAGAACTTACCTGTGCAGTTTGAAGGTGTTCCCTTAAAAATGGTTTTATTCATAGA 101

QY 3611 GACATCAAGCCTGACAAACATACCTCATCGATAGGAGCGGACACATAAAGCTCACCACCTT 3670;

Db 100 GATATTAAACCTGATAATATTTGATTGATCGTGATGTCATATTAAATTTGACTGACTTT 41

QY 3671 GGCCTGTGCACGGGATTCGATGGAGCGACAACTCGAAGT 3710

Db 40 GGCCTCTGCATGGCTTCAGATGGACACACGAGTTCTAAGT 1

RESULT 4

US-09-974-298-152

; Sequence 152, Application US/09974298

; Patent No. US20020156263A1

; GENERAL INFORMATION:

; APPLICANT: Chen, Hui-Mei

; TITLE OF INVENTION: GENES EXPRESSED IN BREAST CANCER

; FILE REFERENCE: PA-0037 P

; CURRENT APPLICATION NUMBER: US/09/974,298

; CURRENT FILING DATE: 2001-10-04

; PRIOR APPLICATION NUMBER: 60/238,331

; PRIOR FILING DATE: 2000-05-10

; NUMBER OF SEQ ID NOS: 194

; SOFTWARE: PERL Program

; SEQ ID NO 152

; LENGTH: 3583

; TYPE: DNA

; ORGANISM: Homo sapiens

; FEATURE:

; NAME/KEY: misc feature

; OTHER INFORMATION: Incyte ID No. US20020156263A1 347975.11

US-09-974-298-152

Query Match 4.1%; Score 232.6; DB 9; Length 3583;
Best Local Similarity 53.1%; Pred. No. 2.2e-41;
Matches 599; Conservative 2; Mismatches 496; Indels 30; Gaps 4;

QY 3108 AGAATCAGCTGAGAGAGATGCACAAAGTGGGACTGCCCATCAGACCCCAAAATCGAGA 3167

Db 417 AAAGAAAGTTAGAAAGGTGATGGAAGAGAGAGGCTTAAAGATGAGGAGAAACGACTCC 476

QY 3168 TCAGGAAATGCTGAACAAAGAGAGAGCACTACATTCGATTGAAGCGCGCAAGATGG 3227

Db 477 GGAGATCAGCATGCTCGGAAGGAAACAGAGTTTCTGTTTGAGAGAAACAGACTTG 536

QY 3228 ACAAGAGCATGTTGTCGTCAACTGAAGCCCATTTGGAGTGGGTGCATTTGGCGAGGTAACGC 3287

Db 537 GATTGGAAGATTTTGAAGTCCGTAAGAGTAATAGGAGAGAGGAGCAATTTGGTGAAGTACGGC 596

QY 3288 TGGTCAGCAAAATCGATACCTCGAACCATTTGTCGATGCGATGAAACCCCTGCGGAAGCGG 3347

Db 597 TTGTTTCAAGAGAGAGATA---CGGACATGTGTATGCAATGAAATATCTCCGTAAGAGC 653

QY 3348 AGTTTCTCAAGCGGAATCAGGTGGCAGACGTCGAAGCGCGAGAGGATATCTCTCGCGGAG 3407

Db 654 ATATGCTTGAAGAAACAGCAGGTTGCGCATCTGTCGCGAGCGTCACATTTCTAGTGGAGG 713

QY 3408 CCGCAATAACTGGTGGTGAAGTTGTACTACAGCTTCCAGGACAGGATATCTGTACT 3467

Db 714 CAGACAGTTTGTGGTGTGAAATGTTCTATAGTTTTCAGGATTAAGCTAAACCTCTACC 773

QY 3468 TTGTCATGACATACACAGTGGTGTGATCTGCTGCTCATCAAACTGGGCATTT 3527

Db 774 TAATCAGAGTTCTGCTGCTGGAGGGGACATGATGACCTTGTGTATGAAAGACACTC 833

QY 3528 TCGAGGAGAACTGCCAGATTCTACATCGCGAGGTCACCTCGCGCGTGGACAGCGTTC 3587

Db 834 TCACAGAGAGGAGACTCAGTTTATATAGCAGAAACAGATATTACCCATAGACTCTATT 893

QY 3588 ACAAAATGGGTTCAATCACAGAGACATCAAGCCTTGACAACATCTCATCGATAGGGAGC 3647

Db 894 ACCAACTTGGATTCTCCACAGAGACATCAAAACAGACAAACCTTCTTTTGGACAGCAAGG 953
Qy 3648 GACACATAAAGCTTACCGACTTTGGCTGTGACCGGATTCGATGACGACCAACACTCGA 3707
Db 954 GCCATGTGAACCTTTCTGACTTTGGTCTTTGGACAGGACTGAAAGAGACATAGGACAG 1013
Qy 3708 AGTACTACAGGAGAGAGGCAATCACTCGCGCCAGGACTCGATGGAGGAAAT 3767
Db 1014 AATTTTA-TAGGAATCTGAACACAGAGCTCCCACTGATTTCACTTTCCAGAACATGAT 1072
Qy 3768 ACTCCGAGAACGAGCAAGCCACCGCTGCTGGAGAGGCGGATGCGGATCAACAAA 3827
Db 1073 TCCAAAGGAACAGCAAAACCTTGGAAAGAAATAGAGTCA----- 1113
Qy 3828 GAGTCCTGGCCCCACTCGCTGGTGGGACCCCGCAACTACATAGCTCCGAGGTGCTGGAGA 3887
Db 1114 ----GCTAGCTTCTCCACAGTAGGCACCTCTGACTACATTTGCTCTGAGGTTCATGC 1169
Qy 3888 GRAGTGGGTACAGCAGCTGTGCGACTACTGGAGCGTGGCGGTCTATCTCTTAYAGATGC 3947
Db 1170 AGACCGGGTACAAAGCTCTGTGATTTGGTGTGCTGGGTGATCATGTATGAGATGC 1229
Qy 3948 TGTGGGTGAGCGGCTTCTGGCAACAGTCCGCTGGAACGCAACAAAGGTCACTCA 4007
Db 1230 TCATCGCTACCCACCTTCTGCTGTGAGACCCCTCAAGACATATAGAGGTGATGA 1289
Qy 4008 ACTGGGAGAAACSTGCTATATTCGCGCGAGCGCGAGTTATCCCGAGGCTACGGACT 4067
Db 1290 ACTGGAAAGAACTTTGACTTTTCTCCAGAAAGTTCCCATCTCTGAGAAAGCAAGGATC 1349
Qy 4068 TGATAAGGAGGCT---CTGTGCGTGGCTGACAGCGGCTGGCAAGAGCTGACAGAGG 4124
Db 1350 TAAATTTGAGGTTCTGCTGTGATGGGACATAGAAATTTGGAGTCTCTGGAGTTGAGGAAA 1409
Qy 4125 TCAAGGACACAGCTTCTTCAAGGACATCGACTTTTGGGACATGCGGAAGCAGAAAGCC 4184
Db 1410 TAAAGTAACCTTTTCTTGAAGCGTGTGACTGGGAACATATCAGAGAGAGACCTGCTG 1469
Qy 4185 CCTACATACCCGGAATCAAGCACCCACGACACATCCAACTTTGAT 4231
Db 1470 CAATATCTATTGAATCAAAAGCATTTGATGATACCTCAAACTTCGAT 1516

RESULT 5

US-09-938-842A-2402
; Sequence 2402; Application US/09938842A
; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE OF INVENTION: SAME, AND METHODS OF USE
; FILE REFERENCE: SCRIP1300-3
; CURRENT APPLICATION NUMBER: US/09/938, 842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: 2001-08-24
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/227, 866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264, 647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300, 111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2402
; LENGTH: 1452
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-938-842A-2402

Query Match 3.5%; Score 200.6; DB 9; Length 1452;
Best Local Similarity 49.6%; Pred No. 1.5e-34;
Matches 640; Conservative 3; Mismatches 617; Indels 31; Gaps 4;

Qy 2984 AAGGAGAGGAGGAGGAGCGCAAGAGGATTCGCGCATCAGGCACTACTCGCCCAAGCCCTTC 3043
Db 25 AACGTCAGGAGGAGGAGGTTCTGGATCGAGCTTGACCATGGAGAAAGTGGCCGACGCG 84
Qy 3044 AAGTCTTTCATGAGCAGCAGCATAGAGAAACGTGATCAAGTCTGATCGCCAGGCACTAT 3103
Db 85 AAGCAGTACATCGAGAATCTACTACAAAGCTCAGAACRAGAACATTTCAAGAGAGAAAGAG 144
Qy 3104 CGCAAGAAATAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCCCGATCAGACCCCAATC 3163
Db 145 AGACGGTGGATCTTAGAAGAAAGTTGGCTCTTCTGGAGTGCCAAAGAGGAGCAAACT 204
Qy 3164 GAGATGAGGAAATGCTGTAACCAAAAGAGAGCAACTACATTCGATTTGAAGCGCGCCAAG 3223
Db 205 AACATGATTAAGAGACCTGGAGAGAAAGAGACAGAGTTTATGAGGCTTAAAGAGCAAG 264
Qy 3224 ATGGACAAGAGCAGTGTTCGTCAAACTGAAGCCCATTTGGAGTGGGTGCTATTTGGGAGGTA 3283
Db 265 ATTAGTCTCGATGACTTTGAGCTTTGACTATCATTTGGAAGAGGTGCTTTTGGTGGAGTT 324
Qy 3284 ACGTGTGTGAGCAAAATCGATACCTCGAAACCATTTGTATGCGATGAAACCCCTCGCGAAA 3343
Db 325 CGCTTATGTCGGGAGAGAAAGTCTGGAA---ATATTTATGCCATGAAGATTTAAGAAA 381
Qy 3344 GCGGACGCTTCTCAAGCGGAATCAGGTGGCACACGTCGAGGCCGAGAGGATATCTCTCGCG 3403
Db 382 TCTGAAATGTCATGAGGACAGGTTGAGCATGTGAGAGCAGAGAGAACCTTGCTGGCT 441
Qy 3404 GAAGCCGACAAATCACTGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAGGATATCTG 3463
Db 442 GAGGTTGAAAGCCATTTATTTGGAAGCTTTACTATTTCATTTCCAGGTCCCAGATCTG 501
Qy 3464 TACTTCTGTGACTACATACAGGTGGTGTGATCTGATGCTGCTCATCAAACTTGGGC 3523
Db 502 TATCTGATTAAGAAATCTCCCGGTGGTGTGATGATGACCTTGCTCATGAGGAGGAG 561
Qy 3524 ATTTTGGAGGAGCACTGGCCGATTTCTACATCGCGAGGTCACTGCGCGCTGGACAGC 3583
Db 562 ACATTACGGGAGACATGTGCCAGATTTTATATGCTCAAGTGTCTCGCCCATTTGAATCC 621
Qy 3584 GTTCACAAAATGGGCTTCAATTCAGAGACATCAAGCCCTGACAACTATCATCATGATAG 3643
Db 622 ATACACAGATACAACTATATTCATAGGATATCAAAACCTGATAACCTACTTTTGGACAA 681
Qy 3644 GACGACACATAAAGCTCAACGACTTTGGCTGTGACCGGATTCGGATGGAGGACACAA 3703
Db 682 GATGGCACATGAACTCTGGACTTTGGGCTCTGTGAAGCCCTCTGTATGTAGAAATTTA 741
Qy 3704 TCGAAGTACTACAGGAGA-----ACGGCAATCACTCGCCGAGG 3743
Db 742 CCTTCAATTCAGGAGAAATAGGCCACGGATGATGAACCTATGTGAGAACCTATGATGTT 801
Qy 3744 ACTCGATGAGCCCTGGGAGGAAATCTCGGAGACGGACCGAACCCCGCTGCTGGAGA 3803
Db 802 GATAGATGCTTTCTCGACACTGATATAACAAGAGAGAGTGGCGAGTCCCGCAGAA--CAAC 859
Qy 3804 GCGAGCGATGCGGATCACCAAGAGTCTCGGCCACTCGCTGGTGGGCGACCCCGAAT 3863
Db 860 TTCAGCAATGGCAGATGATCGCAGAAACTAGCATTTTCACTGTGGGAACACCGGACT 919
Qy 3864 ACATAGCTCCCGAGGTGCTGGAGAGTGGGTATACACGACGTGTGCGACTTACTGGAGCG 3923
Db 920 ATATTGCTCTGAAGTTTGTGCTGAAGAAAGGATATGGCATGGAATGTGATTTGGTGGT 979
Qy 3924 TGGCGTCTATCTTAYGAGATGCTGGTGGTTCAGCCGCCCTTTCTGGCCAAAGTCCCG 3983
Db 980 TAGTGCAATTTATGTACGAAATGCTGTTGGTATPCTCTCTTTTATGCTGATGACCTA 1039
Qy 3984 TGGAAACGCAAAAGGTCTCAACTGGGAGAAACCTGATATTCGCGCGCAGGCGG 4043
Db 1040 TATCAACTGCAAGAAAGATCTCCATTTGGAGAAACCTTTGAATTTTCCGAGGATCGGA 1099

QY 4044 AGTTATCCGGAGGCTACGGACTTGTATAAGAGGCTCTGTGCGTGGCTGACAAAGCGC 4103
 Db 1100 AGTTTTCATCTGAGCAAAAGATCTTATCTGCAGCTTGTGTGCAACGTTGACCATAGC 1159
 QY 4104 TGGCAAGAGCGTGG-----ACGAGGTCAAGAGCCAGCACTTCTTAAGGCGATCGACT 4157
 Db 1160 TTGGTACTTGGAGGAGGAGCCAGCAAAATCAAGGATCATCTTGGTTCAAGGATGTGTGT 1219
 QY 4158 TTGGGACATCGGAAGCAGAAAGCGCCCTACATACCGGAATCAAGCACCRCGAGCA 4217
 Db 1220 GGGAAAGCTCTATGAATGAGGCTGCTGACAAACCAAGTGAACGAGCTAGATA 1279
 QY 4218 CATCAACCTTTGATCCCGTGGATCGGAGAA 4248
 Db 1280 CACAAAATTTTATGAAGTTTGTGAAGTGAA 1310

RESULT 6

US-09-938-842A-1861
 ; Sequence 1861, Application US/09938842A
 ; Patent No. US20020160378A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Harper, Jeff
 ; APPLICANT: Kreps, Joel
 ; APPLICANT: Wang, Xun
 ; APPLICANT: Zhu, Tong
 ; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
 ; TITLE OF INVENTION: SAME, AND METHODS OF USE
 ; FILE REFERENCE: SCRIP1300-3
 ; CURRENT APPLICATION NUMBER: US/09/938,842A
 ; PRIOR FILING DATE: 2001-08-24
 ; PRIOR FILING DATE: 2000-08-24
 ; PRIOR FILING DATE: 2000-06-24
 ; PRIOR FILING DATE: 2001-01-16
 ; PRIOR APPLICATION NUMBER: US 60/227,866
 ; PRIOR APPLICATION NUMBER: US '60/264,647
 ; PRIOR APPLICATION NUMBER: US 60/300,111
 ; PRIOR FILING DATE: 2001-06-22
 ; NUMBER OF SEQ ID NOS: 5379
 ; SEQ ID NO 1861
 ; LENGTH: 1689
 ; TYPE: DNA
 ; ORGANISM: Arabidopsis thaliana
 ; US-09-938-842A-1861

Query Match 3.4%; Score 195.8; DB 9; Length 1689;
 Best Local Similarity 49.7%; Pred. No. 1.9e-33;
 Matches 602; Conservative 4; Mismatches 591; Indels 14; Gaps 4;

QY 3044 AAGTTCTTCATGGAGCAGCAGACATAGAGAACGCTGATCAAGTCTGATCGCCAGCGCACGTAT 3103
 Db 163 AAGCAATATATCGAATCAATTATAAATTCAGAGAAGAGTCTCCAGGAAGAAAGAG 222
 QY 3104 CGAAGATACGCTGGAGAAGAGATGACAAAGTGGAGCTGCCGATCAGACCCAAATC 3163
 Db 223 CGTCGGAGCATCTTGGACAAACACCTAGCTGATGCTGATGTTACTGTGTGAAGACAAGATG 282
 QY 3164 GAGATGAGGAAATCTGAACCAAGAGAGACCACTACATTCGATTGAAGCGCCCAAG 3223
 Db 283 GATATATTAAGAACTTTGAGAAAGGAATGAGTATATGCGTCTCAAGACAGAGAA 342
 QY 3224 ATGGACAAGAGCATGTTCCGTAACACTGAAGCCCATGGAGTGGTGCATTTGGCGAGGTA 3283
 Db 343 ATGGGGTTGATGACTTTGAACTGCTTAGCATCATTTGCCGGGTGCTTTCGGTGAAGTG 402
 QY 3284 ACGTGGTGAGCAAAATCGATACCTCGAACCAATTTGTATGGATGAAGAAACCCCTCGGAAA 3343
 Db 403 AGAATTTGTAAAGAAAATCTACTGAAAGC---GTATATGCAATGAAAAGTTAAAGAAA 459
 QY 3344 GCGGACGTTCTCAACCGGAATCAGGTGCGACAGTGAAGCCGAGAGGATATCTCCGCG 3403
 Db 460 TCCGAGATGCTTCGACGAGGCGAGGTGAACATGTTAAAGCTGAAAGAAATGTGCTTGA 519
 QY 3404 GAAGCCGACATAACTGGTGGTGAAGTTGTACTACAGTCTCCAGGACAAGGATATCTG 3463

Db 520 GAAGTGGATAGCCCTTTTCATTGTCAGCTTTTGTACTCTTTCCAGATGATGAGCATTTG 579
 QY 3464 TACTTTGTATGAGCACTACATACCAGGTGGTATGATGTCGCTCTCATCAAACTGGC 3523
 Db 580 TATCTTATATGGAATACCTCCCTGGAGTGATATGATGACACTGCTGTGGAAGAGT 639
 QY 3524 ATTTTCGAGGAGAACTGGCCAGATTCTACATCGCCGAGGTCACTCGCCGCTGGACAGC 3583
 Db 640 ACCTTACGGAGATGAGACTCGGTTTTATGTTGGCGAGACAATTTCTGGCTATTGAGTCT 699
 QY 3584 GTTCAAAATATGGGCTTCAATTCACAGAGACATCAAGCCCTGACAACTACTCATCATAGG 3643
 Db 700 ATCCATAAGCAATAATTACGTCCACAGGATATAAAGCCTGATAAATTTATTGATTACTGA 759
 QY 3644 GACGACACATAAAGCTCACCGACTTGGCCCTGTCACGGGATTCGGATGGACGACACAAC 3703
 Db 760 AACGCCATATCAAGCTTTTCAGATTTTGGATTGACAAAGTCTCTGGAAGCAAAAATTTT 819
 QY 3704 TCGAAGTACTACCAAGG-----GAACGGCAATCACTCGCCGAGGACTCGATGGAGCC 3756
 Db 820 CCAGATTTTAAGCGGAGCTTGTTCACAGGAGTACAAGCCCTGACAGACAATCATAGA 879
 QY 3757 CTGGAGGAATACTCCGAGAACGGACCCACCGTCTGTGGAGAGCGGACGATCGC 3816
 Db 880 CTCTCAAGCCCTCTTCTGCACCTAGGAGAACTCAGCA-GGAACAGCTTTTACATTGGCA 938
 QY 3817 CGATCACCAGAGAGTCTTGGCCCACTCGTGTGGGACCCGACCACTACATAGTCCCGA 3876
 Db 939 ACAAAACAGAGGACCCCTGGCTTTTCTACAGTAGGAACTCCCGATTACATTTGCCCTGA 998
 QY 3877 GGTGCTGGAGAGTGTGTACACGACGCTGTGCCACTACTCGAGCGTGGGCGTCTCATCT 3936
 Db 999 GGTGCTGTGAAGAAAGGTATGGAATGAGTGTGATTTGGTGTCTCTTGGAGCAATCAT 1058
 QY 3937 YTAGAGATGCTGTGGTTCAGCCGCCCTTTCTGCCCACAGTTCGCTGGAAACCAACA 3996
 Db 1059 GTTCGAGATGCTGTAGGTTTCCGCCATTTCTATCCGAAGAACCTTTTGGCAACATGTAG 1118
 QY 3997 AAGGTCTACACTCGGAGAAACSCCTGCATATTCGCGCGAGCGGAGTATCCCGCGA 4056
 Db 1119 AAAGATTCTAAACTCGAAACCTGCTTGAAGTTCCTGATGAAGCTTAAGCTCTCCATCGA 1178
 QY 4057 GGCTACGACTTGATAAGGAGGCTCTGTGCGTCTGACAGCGGCTGGG---CAAGAG 4113
 Db 1179 GGTAAAGATCTTATCCGAAGACTGCTCTGCATGTCGAACAGAGGCTTGGACCAAGG 1238
 QY 4114 CGTGACGAGGTCAAGACCCACGACTTCTTCAAGGGCATCGACTTTGGGAGCATCGGAA 4173
 Db 1239 AGTTCAAGAAATCAAGACACACCCCTTGGTTTAGGGAGTGAATGGGAACGGCTATATGA 1298
 QY 4174 CGAGAAAGCGCTACATACCGGAATCAAGCACCCRCAGGACACATCAACTTTGATCC 4233
 Db 1299 GTCAATGCTCCATATATACCAACAAGTGAAGCATGAGCTTGATACCAAAACTTTGAAA 1358
 QY 4234 CGTGATCCGG-4244
 Db 1359 GTTGTATGAGG 1369

RESULT 7

US-09-771-161A-89
 ; Sequence 89, Application US/09771161A
 ; Patent No. US2002011081A1
 ; GENERAL INFORMATION:
 ; APPLICANT: LEVINE, et al.
 ; TITLE OF INVENTION: VARIANTS OF PROTEIN KINASES
 ; FILE REFERENCE: 802620-2005.1
 ; CURRENT APPLICATION NUMBER: US/09/771,161A
 ; CURRENT FILING DATE: 2001-01-26
 ; PRIOR APPLICATION NUMBER: 09/724,676
 ; PRIOR FILING DATE: 2000-11-28
 ; PRIOR APPLICATION NUMBER: 136776

Fri Jan 17 11:17:11 2003

PRIOR FILING DATE: 2000-06-15
PRIOR APPLICATION NUMBER: 135619
PRIOR FILING DATE: 2000-04-12
NUMBER OF SEQ ID NOS: 273
SOFTWARE: PatentIn version 3.0
SEQ ID NO 89
LENGTH: 1818
TYPE: DNA
ORGANISM: Homo sapiens
US-09-771-161A-89

Query Match 3.3%; Score 188.6; DB 10; Length 1818;
Best Local Similarity 54.9%; Pred. No. 7.4e-32;
Matches 469; Conservative 2; Mismatches 356; Indels 28; Gaps 4;
QY 3108 AGAATCAGCTGGAGGAGAGATGACAAAGTGGGACCTGCCGATCAGACCCCAATCGAGA 3167
DB 729 AAAAGAACTTAGAAAAGCTGATGGAAGAAGAGCGCTTAAAGATGAGGAGAAACGACTCC 788
QY 3168 TCAGGAAATGCTGAACCAAAAGAGAGCACTACATTCGATTGAAGCGGCCCAAGATGG 3227
DB 789 GGAGATCAGCAGATGCTCGGAGAGAAACAGAGTTTCTCGTTGAAGAGAAACAGACTTG 848
QY 3228 ACAAGAGCATGTTCTGCAAACTGAAGCCCATTTGAGTGGTGCATTTGGCGAGGTACGC 3287
DB 849 GATTGGAAGATTTTCAGTCTTAAAGTAAATAGCAGAGGAGCATTTGGTGGAGTACGCC 908
QY 3288 TGGTGAGCAAAATCATCTCGACCTCGAACCATTTGTATGCGATGAAACCCCTCGGAAACGG 3347
DB 909 TTGTTTCAAGAAGAAGATA---CGGACATGTTGATGCAATGAAATACTCGTAAAGCAG 965
QY 3348 ACCTTCTCAAGCGGAATCAGTGGGACACAGCTGAAGCGCGAGGAGGATATCTCCGGAAG 3407
DB 966 ATATGCTTGAAGAAGAGCAGTTGGCCACATTCGTGCGGAGCGTGACATTTAGTGGAGG 1025
QY 3408 CCGACAAATCACTGGGTGGTGAAGTTGTACTACAGCTTCCAGGACAAAGATATCTGTACT 3467
DB 1026 CAGACAGATTTTGGGTGTTGAAATGTTCTATAGTTTTCAGGTAAGCTAAACCTTACC 1085
QY 3468 TTGTGATGAGTACATACACAGTGGTGTGATCTGATGCGCTGATCAATCAAACTGGGCATTT 3527
DB 1086 TAATCATGAGTTCTYTCCTCGGAGGAGCATGATGACCTTGTGTATGAAAGAAAGACATC 1145
QY 3528 TCGAGGAGGAACCTGCGCAGATTTCTACATCGCGAGTCCACCTGCGCGTGGACAGCGTTC 3587
DB 1146 T-GACGAGAGAGAGACTCAGTTTATATAGCAGAAACAGTATTAGCCATAGACTCTATTC 1204
QY 3588 ACAAAATGGGCTTCATTCACAGAGACATCAAGCCTGACAAACATATCTCATCTGATAGGACG 3647
DB 1205 ACCAATTTGGATTATCCACAGAGACATCAACACAGACACCTTCTTTTGGACAGCAAGG 1264
QY 3648 GACACATAAAGCTCAGCAGCTTTGGCTGTGACGCGGATTCGATGGAGCGCCCTGGAGGAAT 3767
DB 1265 GCCATGTAAATTTCTACATTTTGGCTTTTGCACAGAGCTGAAAGAACACATAGGACAG 1324
QY 3768 AGTACTACAGGAGCAAGCGCAATCAGTTCGCGGAGGAGTCTGATGGAGCGCTGGAGGAAT 3767
DB 1325 AATTTTA-TAGGATCTGAACACAGCTTCCCGAGTTCATTTTCCAGAAACATGAAT 1383
QY 3768 ACTCCGAGAACCGAGCCCAACCGTGTGGAGAGCGGACGAGTTCGCGGATCAGCAAAA 3827
DB 1384 TCCAAAGGAAAGCAGAAACCTTGAAGAAGAAATAGAGCTGAG----- 1425
QY 3828 GAGTCTGCGCCACTCGCTGTGGGACCCCGAACTACATAGCTTCCCGAGGTCTGGAGA 3887
DB 1426 -----CTAGCTTCTCCACAGTAGGACCTCTGACTACATTTGCTTCCCTGAGGTGTTATGC 1480
QY 3888 GRAGTGGGTACACGAGCTGTGCGACTACTGAGCGGTGGGCTCATCTTAYGAGATGC 3947
DB 1481 AGACCGGGTACAAACAGCTCTGTGATTTGGTGGTGGCTGATCATGTATGAGATGC 1540
QY 3948 TGGTGGGTGAGCCGC 3962

Db 1541 TCATCGGTAAAGTTGC 1555

RESULT 8

US-09-879-536-66
Sequence 66, Application US/09879536
Patent No. US20020144298A1
GENERAL INFORMATION:
APPLICANT: Endege, Wilson O.
APPLICANT: Steinmann, Kathleen E.
APPLICANT: Astle, Jon H.
APPLICANT: Burgess, Christopher C.
APPLICANT: Bushnell, Steven E.
APPLICANT: Carroll III, Eddie
APPLICANT: Catino, Theodore J.
APPLICANT: Derti, Adnan
APPLICANT: Ford, Donna M.
APPLICANT: Lewis, Marcia E.
APPLICANT: Monahan, John E.
APPLICANT: Schlegel, Robert
TITLE OF INVENTION: NOVEL HUMAN GENES AND GENE EXPRESSION
TITLE OF INVENTION: PRODUCTS
FILE REFERENCE: CCD-257 (US)
CURRENT APPLICATION NUMBER: US/09/879,536
CURRENT FILING DATE: 2001-09-21
PRIOR APPLICATION NUMBER: US 60/088,801
PRIOR FILING DATE: 1998-06-10
NUMBER OF SEQ ID NOS: 850
SOFTWARE: FastSeq for Windows Version 3.0
SEQ ID NO 66
LENGTH: 678
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc.feature
LOCATION: (1)-(678)
OTHER INFORMATION: n = A,T,C or G
US-09-879-536-66

Query Match 2.9%; Score 164.6; DB 10; Length 678;
Best Local Similarity 63.2%; Pred. No. 8.4e-27;
Matches 285; Conservative 0; Mismatches 161; Indels 5; Gaps 2;

QY 3078 TCAGTCTGATCGCCGACGATATCGGAAAGATCAGTGGAGAGGAGATGCACAAAG 3137
DB 3 TCAATCTCATCAGCAGCTGTACATCGTAAAAACAATAGAGAATGAATGATCGCGG 62
QY 3138 TGGGACTGCCCGATCAGACCCAAATCGAGATGAGGAAATGCTGAACCAAGAGGAGCA 3197
DB 63 TTGGATTATCTCAAGATGCCGAGGATCAAAATGAGAAAGATGCTTTGCCAAAGATCTA 122
QY 3198 ACTACATTCGATTGAAGCGGCCCAAGATGGACAGAGCATGTTTCGCAAACTGAAGCCCA 3257
DB 123 ATTACATCCGCTTTAAAGGGCTTAAATGGACAGTCTATGTTTGAAGATTAAGACAC 182
QY 3258 TTGGAGTGGGTGCTATTTGGGAGGTAAACGCTGGTGACCAAAATCGATACCTCGAACCA 3317
DB 183 TAGGAATAGGAGCATTTGGTGAAGTCTGTCTAGCAAGAAAGTAGATA---CTAAGCCT 239
QY 3318 TGTATGCCATGAARACCTCGGAAAGCGGACGTTCTCAAGCGGAATCAGGTGGCACAG 3377
DB 240 TGTATGCCAAGAAACTCTTCGAAAGAAGATGTTCTTCTCGAAATCAAGTCGCTCATG 299
QY 3378 TGAAGCGCGAGAGGATATCTTCGCGGAGCGGACATTAAGTGGTGTGAAGTTGACT 3437
DB 300 TTAAGCTGAGAGATATCTTGGCTGAAGCTGAATGAATGGGTAGTTCGCTATATT 359
QY 3438 ACAGCTTCCAGCAAGGATATCTGTTGATGGACTACATACAGGTGGTGATC 3497
DB 360 ATTCAATCCAGATAAGGCCCATTTATCTTT--GTAATGGCTACATCTCTTNGGCTGATA 417
QY 3498 TGATGCTGCTGCTCATCAAACTGGGCATTTT 3528

Db 662 TGAAACACCTCAAGAGACATACAGAAAGTGATGAACCTGGAAAGAACTCTGGTATTTC 721

; Sequence 2503, Application US/09938842A

```

; Patent No. US20020160378A1
; GENERAL INFORMATION:
; APPLICANT: Harper, Jeff
; APPLICANT: Kreps, Joel
; APPLICANT: Wang, Xun
; APPLICANT: Zhu, Tong
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING
; FILE REFERENCE: SAME, AND METHODS OF USE
; CURRENT APPLICATION NUMBER: US/09/938,842A
; CURRENT FILING DATE: 2001-08-24
; PRIOR APPLICATION NUMBER: US 60/227,866
; PRIOR FILING DATE: 2000-08-24
; PRIOR APPLICATION NUMBER: US 60/264,647
; PRIOR FILING DATE: 2001-01-16
; PRIOR APPLICATION NUMBER: US 60/300,111
; PRIOR FILING DATE: 2001-06-22
; NUMBER OF SEQ ID NOS: 5379
; SEQ ID NO 2503
; LENGTH: 1416
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
; US-09-938-842A-2503

Query Match      2.4%; Score 138.6; DB 9; Length 1416;
Best Local Similarity 58.7%; Pred. No. 6.4e-21;
Matches 259; Conservative 0; Mismatches 179; Indels 3; Gaps 1;

QY 3247 ACTGAAGCCATTTGGAGTGGTGCATTTGGCGAGGTAAACCTGGTGAAGCAAAATCGATAC 3306
Db 426 ATTGAAGTTTGTGGCAAGAGTGCATTTGGAAAGTGTACCAAGTGGAGGAAAGACAC 485
QY 3307 CTCGAACCATTTGTATCGGATGAAACCCCTGCGGAAAGCGGAGTCTCTCAAGCGGAATCA 3366
Db 486 GTC---TGATATACCGATGAAGTTCATGAGAAAGATAAATTTTGAGAGATCA 542
QY 3367 GGTGGCAGACGTGAAGCCCGAGAGGATATCTCTCGGGAAGCCGACAAATAACTGGGTGT 3426
Db 543 TGTGAATACATGAAAGCCCGAGGCGGATATCTTAACCAAAATCGATCATCTTTTCATTGT 602
QY 3427 GAAGTTGTACTACAGCTTCAGGACAGGATATCTGTACTTTGTGTGATGACTACATACC 3486
Db 603 GCAACTTAATATCTTTTCAGACCAAAATACAGATTGTATCTTTGACTTTATAAA 662
QY 3487 AGTGTGTATCTGTATGCTGTCTATCAAACTGGGCAATTTTCGAGGAGAACTGGCCAG 3546
Db 663 CGCAGGTGATCTTTTCTCCAGCTATCATCAAGGCTTTTCAGGAGGACTTGGCTCG 722
QY 3547 ATTCTACATCGCCGAGGTACCTGCGCGGTGGACAGGTTTCACAAATAGGCTTCATCA 3606
Db 723 TGTGTACACTGCAGAAATCTCTCTGCAGTTTCCCATCTCCATGAGAAAGGCATAATGCA 782
QY 3607 CAGAGACATCAAGCCGTGACAACTACTCATCGATAGGCGGACACATAAAGCTCACCGA 3666
Db 783 TAGAGATCTTAAACCTGAAACATACTCATGGACGTAGATGGCCATGTGTATGTTAACTGA 842
QY 3667 CTTTGGCTGTGCACGGGATT 3687
Db 843 TTTTGGTTTAGCAAGAATTT 863

RESULT 12
US-09-880-107-2146
; Sequence 2146, Application US/09880107
; Patent No. US20020142981A1
; GENERAL INFORMATION:
; APPLICANT: Horne, Darci T.
; APPLICANT: Vockley, Joseph G.
; APPLICANT: Scherf, Uwe
; APPLICANT: Gene Logic, Inc.
; TITLE OF INVENTION: Gene Expression Profiles in Liver Cancer
; FILE REFERENCE: 44921-5028-WO
; CURRENT APPLICATION NUMBER: US/09/880,107

```

```

; CURRENT FILING DATE: 2001-06-14
; PRIOR APPLICATION NUMBER: US 60/211,379
; PRIOR FILING DATE: 2000-06-14
; PRIOR APPLICATION NUMBER: US 60/237,054
; PRIOR FILING DATE: 2000-10-02
; NUMBER OF SEQ ID NOS: 3950
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 2146
; LENGTH: 3061
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; OTHER INFORMATION: Genbank Accession No. US20020142981A1 L07597
; US-09-880-107-2146

Query Match      2.4%; Score 134.6; DB 10; Length 3061;
Best Local Similarity 58.0%; Pred. No. 7.3e-20;
Matches 258; Conservative 0; Mismatches 184; Indels 3; Gaps 1;

QY 3239 TTCGTCAAACTGAAGCCCATTTGGTGGTGCATTTGGCGAGGTAAAGCTGGTGAGCAAA 3298
Db 230 TTCGAGCTCTCAAGTTCTTGGCCAGGATCTTTGGCAAAAGTCTTCTCGTGGTGGGAAA 289
QY 3299 ATCGATACCTCGAACCATTGTTGATGCG---ATGAAAACCCCTGCGGAAAGCGGACGTTCTC 3355
Db 290 GTCACCGCGGCTGCACAGTGGCCACCTGTATGCTATGAAGTGTCTGAAGAAGCAACGCTG 349
QY 3356 AAGCGGAAATCAGTGGCACACAGCTGAAGCCGAGAGGATATCTTCGCGGAAAGCGGCAAT 3415
Db 350 AAAGTACGTGACCGCGCTCCGGAACCAAGATGGAGAGACATCTCTGCTGATGATAATCAC 409
QY 3416 AACTGGTGTGTAAGTTGTACTACAGCTTCAGGACAAAGGATAATCTGTACTTTGTGATG 3475
Db 410 CCATCTGCTGTGAAGTGTCACTATGCTTCCAGCCGAGGCAAGCTCTATCTCATCTCTG 469
QY 3476 GACTACATACAGGTGGTGTATGCTGCTGCTCATCAAACTGGGCAATTTTCGAGGAG 3535
Db 470 GACTTCTGCTGTGGGACCTCTTCAACCGGCTCTCAAAAGAGGTGATGTTTCACGGAG 529
QY 3536 GAATGSCCAAGATCTACATCGCGGAGGTACCTGGCGGTGACAGCGTTCACAAATG 3595
Db 530 GAGATGTGAAGTTTACCTGGCGGAGCTGGCTCTGGGCGCTGGATCACCCTGCACAGCTG 589
QY 3596 GGCTTCATTCACAGACATCAAGCTGCACACATATCATCATAGGAGGACGACACATA 3655
Db 590 GGTATCTTACAGAGACCTCAAGCCTGAGAACATCTTCTGATGAGGAGGCGCCACATC 649
QY 3656 AAGCTCACGACTTTGGCCTGTGCA 3680
Db 650 AAACCTCACTGACTTTGGCCTGAGCA 674

RESULT 13
US-09-971-845-1
; Sequence 1, Application US/09971845
; Patent No. US20020132247A1
; GENERAL INFORMATION:
; APPLICANT: DELANEY, ALLEN
; APPLICANT: YOGANATHAN, THILLAINATHAN
; TITLE OF INVENTION: DYSTROPHIA MYOTONICA PROTEIN KINASE
; FILE REFERENCE: (DM-PK) AND ITS USES
; FILE REFERENCE: KINE028CON
; CURRENT APPLICATION NUMBER: US/09/971,845
; CURRENT FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: PCT/US01/21479
; PRIOR FILING DATE: 2001-09-20
; PRIOR APPLICATION NUMBER: 60/238,558
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 1
; LENGTH: 3407
; TYPE: DNA

```


OTHER INFORMATION: Genbank Accession No. U520020142981A1 M34182
US-09-880-107-2340

Query Match 2.2%; Score 127; DB 10; Length 1635;
Best Local Similarity 56.8%; Pred. No. 2.5e-18;
Matches 254; Conservative 0; Mismatches 190; Indels 3; Gaps 1;

QY 3238 GTTCGCTCAAACTGAAGCCCAATTGGAGTGGGTGCAATTGGCGAGGTAACGCTGGTGAGCAA 3297
DB 181 GTTCGAACGGCTCAGGACGCTGGGCATGGGCTCCTTCGGCGGGGTGATGCTGGTGAGGCA 240
QY 3298 AATCGATACCTCGAACCAATTGTTATGGATGAACCCCTGCGGAAGCGGACGTTCTCAA 3357
DB 241 CCAGGACACGGGGCCA---CTACGCCATGAAGATCCTCAACAGCAGAAGGTGGTGAA 297
QY 3358 GCGGAATCAGGTGGCACACGTGAAGCGGAGAGGGATATCCTCGCGGAAGCCGACAATAA 3417
DB 298 GATGAAGCAGGTGAGCACATCTGAACGAGAGGCCATCCTGCGAGCGATCGACTTTCC 357
QY 3418 CTGGGTGGTGAAGTTCTACTACAGCTTCCAGGACAGGATAATCTGTACTTTGTGATGGA 3477
DB 358 GTTCCTGGTCAAGCTCCAGTTCTCTTTAAGGACAACTCCTACCTGTACCTGGTGATGGA 417
QY 3478 CTATATACCAGGTGGTGATCTGATGCTGCTCATCAAACTGGGCATTTTCGAGGAGGA 3537
DB 418 GTACGTGCGGGTGGGAGATGTTCTCCGCCCTACAGCGCGTGGGAAGTTTAGCGAGCC 477
QY 3538 ACTGGCCAGATTCTACATCGCGAGGTACCTGCGCCGTGGACACGCTTCACAAAATGGG 3597
DB 478 CCATGCGCTTTCTATGCGCCAGGTGCTCTGCGCGTCCAGTACCTACACTCGCTGA 537
QY 3598 CTTTCATTCACAGACATCAAGCCTGACACATATCTATGAGGACGAGCACATAAA 3657
DB 538 CCTCATCCACCGGACCTGAGCCCGGAGAACTCTCTCATCGACGACGAGGCTACCTGCA 597
QY 3658 GCTACCGACTTTGGCCTGTGACGCGG 3684
DB 598 GGTGACGACTTCGGTTTCGCCAAGCG 624

Search completed: January 16, 2003, 22:07:03
Job time : 303.731 secs

Result No.	Score	Query	Length	DB	ID	Description
c						
1	890.8	15.6	1101	17	CNS0145B	AT104549 Drosophil
2	643.6	11.3	662	13	BI637464	BI637464 SD19495.5
3	340.8	6.0	930	14	BQ980648	BQ980648 AGENCOURT
4	334.8	5.9	737	14	BM950668	BM950668 UI-M-EHOP
5	320.6	5.6	644	9	A1517339	A1517339.GH28053.5
6	304	5.3	689	13	BI331257	BI331257.602982912

QY	2123	CCCAACTGTGTCSCCT	2139
Db	50	CCCATAMTGTGWACNT	34
RESULT 2			
BI637464			
LOCUS			
DEFINITION			
<p>BI637464 662 bp mRNA linear. EST 10-SEP-2001 SD19495.5.prime SD Drosophila melanogaster Schneider L2 cell culture pot2 Drosophila melanogaster cDNA clone SD19495 5 similar to wts: Fban0012072 GO:[protein phosphorylation (GO:0006468)]; protein serine/threonine kinase (GO:0004674); eye morphogenesis (GO:0007456); protein serine/threonine kinase (GO:0004674); protein kinase (GO:0004672); imaginal disc growt, mRNA sequence.</p>			
ACCESSION			
VERSION			
KEYWORDS			
SOURCE			
<p>BI637464 1 GI:15539674 EST. fruit fly. Drosophila melanogaster Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Ephydroidea; Drosophilidae; Drosophila. 1 (bases 1 to 662) Harvey,D., Brokstein,P., Hong,L., Evans-Holm,M., Su,C., Tsang,G., Lewis,S. and Rubin,G.M. BDGP/BHMI Drosophila EST Project Unpublished (2001) Contact: Stapleton, M. BDGP</p>			
REFERENCE			
AUTHORS			
TITLE			
JOURNAL			
COMMENT			
<p>One Lawrence Berkeley National Lab Cyclotron Rd, Berkeley, CA 94720, USA Fax: 510 486 6798 Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu hit genomic AE003775:-arm:3R [26377421,26613529] estimated-cyto:100A3-100B2: 05/19/2001 Plate: SD.194 row: H column: 11 High quality sequence stop: 604. Location/Qualifiers 1..662 /oranism="Drosophila melanogaster" /db_xref="taxon:7227" /clonae="SD19495" /cclone.lib="SD Drosophila melanogaster Schneider L2 cell culture pot2" /lab_host="DH5-alpha" /note="Vector: pOT2; Site:1: ECORI; Site 2: XhoI; Sized fractionated cDNAs were directly ligated into pOT2." Plasmid cDNA library. 154 a 200 c 190 g 118 t</p>			
BASE COUNT			
ORIGIN			
<p>Query Match 11.3%; Score 643.6; DB 13; Length 662; Best Local Similarity 99.2%; Pred No. 5.2e-103; Matches 657; Conservative 1; Mismatches 0; Indels 4; Gaps 1</p>			
QY	955	CAGAGCCAATCAAATTCCCGGCCGGATGTGCCCCAGTGTTGTGC-----GTGTGTGT	1010
Db	1	CAGAGCCAATCAAATTCCCGGCCGGATGTGCCCCAGTGTTGTGTGTGTGTGTGT	60
QY	1011	GTTGTGTGCTGTGCTGTGCGAGTCTTAGTGTGCGAGCATTTCTGTGTATATGAG	1070
Db	61	GTTGTGTGCTGTGCTGTGCTGTGCGAGTCTTAGTGTGCGAGCATTTCTGTGTATGAG	120
QY	1071	TGCTAAATGCCAGGGCGGAAGCAGCAGCATCATGATCCAGCGGGCGAAAAAGGGGGC	1130
Db	121	TGCTAAATGCCAGGGCGGAAGCAGCAGCATCATGATCCAGCGGGCGAAAAAGGGGGC	180
QY	1131	GTGCCCCCAATGATAAATACACGGCGAACCCCTCGAGACATCACAGCAGCCTTAACC	1190
Db	181	GTGCCCCCAATGATAAATACACGGCGAACCCCTCGAGACATCACAGCAGCCTTAACC	240
QY	1191	GATTGTAAGTACAAAATAAACCATAGGAATAATACAGAATTACACACCTCTCGCATACCG	1250

Db 241 GATTTGAAGTACAAAATAACCATAGGAATATACAGAAATACACACTCTGCGATACACGG 300
 QY 1251 CGACCAACGGGACGACGATGACCTACTCTGACTATACACACGCGGACGAGCGATGG 1310
 Db 301 CGACCAACGGGACGACGATGACCTACTCTGACTATACACACGCGGACGAGCGATGG 360
 QY 1311 AGCCGCCACCTCCGCTCTCTGCTCGGAGCTGGTATACGCGCGCGCGCGCATTTG 1370
 Db 361 AGCCGCCACCTCCGCTCTCTGCTCGGAGCTGGTATACGCGCGCGCGCGCATTTG 420
 QY 1371 TAGGTACGCGGACGCGGCTCCATATCCGATCCGCTGCGGCTGGGCTGGGCTGG 1430
 Db 421 TAGGTACGCGGACGCGGCTCCATATCCGATCCGCTGCGGCTGGGCTGGGCTGG 480
 QY 1431 TGGCCAAACGGAGCTGSCCAAGATGATCAGCGCCCTTAATGCCAAACAACATGATCCGGA 1490
 Db 481 TGGCCAAACGGAGCTGSCCAAGATGATGACGGCCCTTAATGCCAAACAACATGATCCGGA 540
 QY 1491 AGCCGAGCATCGAAGCGGACGCGGAGCAGTCACTACTGCGCTGCACTCGGCTCTGG 1550
 Db 541 AGCCGAGCATCGAAGCGGACGCGGAGCAGTCACTACTGCGCTGCACTCGGCTCTGG 600
 QY 1551 ACTCGGAGCGGCTAGCTCCGATCGGACAGCGCGGAGCAGTCACTACTGCGCTGCACTCGGCTCTGG 1610
 Db 601 ACTCGGAGCGGCTAGCTCCGATCGGACAGCGCGGAGCAGTCACTACTGCGCTGCACTCGGCTCTGG 660
 QY 1611 GC 1612
 Db 661 GC 662
 RESULT 3
 BQ898648
 LOCUS
 DEFINITION AGENCOURT_8489599 Lupski_dorsal_root_ganglion Homo sapiens cDNA
 clone IMAGE:6180847-5', mRNA sequence.
 ACCESSION BQ898648
 VERSION BQ898648.1 GI:22290662
 KEYWORDS EST.
 SOURCE human.
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
 REFERENCE 1 (bases 1 to 930)
 AUTHORS NIH-MGC <http://mgc.nci.nih.gov/>.
 TITLE National Institutes of Health, Mammalian Gene Collection (MGC)
 JOURNAL Unpublished (1999)
 COMMENT Contact: Robert Strausberg, Ph.D.
 Email: cgapbs-r@mail.nih.gov
 Tissue Procurement: Dr. James R. Lupski
 cDNA Library Preparation: Life Technologies, Inc.
 DNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
 Distribution by: Agencourt Bioscience Corporation
 Cloned through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
 Note: LLNL13564
 High quality sequence stop: 637.
 Location/Qualifiers
 1. 930
 /organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone="IMAGE:6180847"
 /clone_lib="Lupski_dorsal_root_ganglion"
 /sex="male"
 /tissue_type="dorsal root ganglia"
 /dev_stage="adult, 36 yr"
 /lab_host="DH10B"
 /note="vector: pCMV-SPORT6 (Life Technologies); Site_1:
 Note; Site_2: SalI; cDNA made by oligo-dT priming.
 Directionally cloned using the following adaptors:
 5'-TCGACCCACCGGCTCCG-3' and

FEATURES source

5'-GACTAGTTCTAGATCGGAGCGCGCCT(15)-3'. Size selected >
 1 kb for average insert length 1.7 kb. This is a primary
 library, non-amplified. Library constructed by Life
 Technologies and donated by J. Lupski, M.D./Ph.D. (Baylor
 College of Medicine) and is available through Life
 Technologies.
 BASE COUNT 266 a 229 c 261 g 174 t
 ORIGIN
 Query Match 6.0%; Score 340.8; DB 14; Length 930;
 Best Local Similarity 70.2%; Pred. No. 6.4e-50;
 Matches 486; Conservative 0; Mismatches 202; Indels 4; Gaps 2;
 QY 2993 GAGGAGGAGCGCAAGAGTTCGCGCATCAGGAGTACTCGCGCAAGCTTCAAGTTCTTC 3052
 Db 93 GACGAGAGAGAGAGAGTACGCGATCAAGAGCTACTCGCCATAGCCCTTTAAGTTCTTC 152
 QY 3053 ATGGAGCAGCAGATAGAGAACGTGATCAAGTGTGTATCGCCAGCGCAGTATCGCAAGAT 3112
 Db 153 ATGGAGCAGCAGTGGAGAAATGTCAAAACCTACCAGCAGAGGTTAACCCGAGGCTG 212
 QY 3113 CAGCTGGAGAGGAGATGCACAAAGTGGGACTGCCGATCAGACCCAAATCGAGATGAGG 3172
 Db 213 CAGCTGGAGAGAAATGCCAAAGCTGGACTCTGTGAAGCTGACGAGGAGCAGATGCGG 272
 QY 3173 AAAATGCTGAACCAAGAGAGCAACTACATTCGATTGAAGCGCGCCAAAGATGACCAAG 3232
 Db 273 AGATCTCTACCAGAAAGAGTCTAATTACACAGGTTAAGAGGGCCCAAGATGACCAAG 332
 QY 3233 AGCATGTTCTGCTCAAACTGAAGCCCATTTGAGTGGTGGTATTTGGCGAGGTACACGCTGGT 3292
 Db 333 TCTATGTTTGTCAAGATCAAAACCTCGGATCGGTCTTTGGAGAAAGTGTGCTTCT 392
 QY 3293 AGCAAAATCGATACCTCGAACCATTTGTATGGATGAACACCTCCGGAAGCGGACGTT 3352
 Db 393 TGTAAAGTGGG--CACTCAGCGCTGTACGCCATGAAGACCTTAAGGAAAAGATGTC 449
 QY 3353 CTCAGCGGAAATCAGTGGCAGCAGCTGAAGCGGAGAGGATATCTCGCGAAGCCGAC 3412
 Db 450 CTGAACCGGAAATCAGTGGCAGCAGCTGAAGCGGAGAGGATATCTCGCGGAGGACAC 509
 QY 3413 AATAACTGGTGGTGAAGTGTGTACTACAGTTCCAGGACAAAGGATTAATCTGTACTTTGTG 3472
 Db 510 AATGATGGTGGTGTCAAACTCTACTACTCTTCCAAAGACAGACAGCTGTACTTTGTG 569
 QY 3473 ATGGACTACACAGGTTGGT 3532
 Db 570 ATGGACTACATCCCTGGTGGGAGCATGATGAGCTGTGTGTGTGTGTGTGTGTGTGTGT 629
 QY 3533 GAGGAACCTGGCAGATTTCTACATCGCGAGGTCACTCGCGGTGACAGGCTTCCACAAA 3592
 Db 630 GAGCAGCTGGCGGCTTCTACATCGCAGAGCTGACTTTGGCCATGACAGTGTCCACAG 689
 QY 3593 ATGGCTTCAATCAGAGACATCAAGCTGACACATCTATCATGATAGGAGGAGGACAC 3652
 Db 690 ATGGCTTCAATCAGAGACATCAAGCTGACACATCTATCATGATAGGAGTGTGATGTC-C 748
 QY 3653 ATAAAGCTCACCGACTTTGGCTGTGCACGGG 3684
 Db 749 ATAAAGCTCACAGATTTCCGGCTCTGCACGTG 780
 RESULT 4
 BQ950668
 LOCUS
 DEFINITION UI-M-EH0p-buu-b-12-0-UI.r1 NIH_BMAP_EH0p Mus musculus cDNA clone
 IMAGE:5686955 5', mRNA sequence.
 ACCESSION BQ950668
 VERSION BQ950668.1 GI:19434258
 KEYWORDS EST.
 SOURCE house mouse.
 ORGANISM Mus musculus
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
1 (bases 1 to 737)
NIH-MGC <http://mgc.nci.nih.gov/>
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cdapbs-r@mail.nih.gov
Tissue Procurement: Dr. James Lin, University of Iowa
cDNA Library preparation: Dr. M. Bento Soares, University of Iowa
cDNA Library Arrayed by: Dr. M. Bento Soares, University of Iowa
DNA Sequencing by: Dr. M. Bento Soares, University of Iowa
Clone Distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LNL at:
<http://image.lnl.gov>
This clone was contributed by the Brain Molecular Anatomy Project
(BMAP)

FEATURES	SOURCE
----------	--------

```

I. : 737
source
/organism="Mus musculus"
/strain="C57BL/6"
/db_xref="taxon:10090"
/clone="IMAGE:568695"
/clone_lib="NIH_BMAP_EH0p"
/tissue_type="whole brain"
/dev_stage="embryo 18.5 dpc"
/lab_host="DH10B (T1 phage resistant)"
/site_1: Ecor I;
/site_2: Not I; The library was constructed according to
Bonaldo, Lennon and Soares, Genome Research, 6:791-806,
1996. Denatured mRNA was size fractionated on a 1% agarose
gel. First strand cDNA synthesis was primed with an
oligo-dT primer containing a Not I site. Double stranded
cDNA was size selected according to mRNA size fraction,
ligated with Ecor I adaptor, digested with Not I, and then
cloned directionally into pYX-Asc vector. The library tag
sequence located between the Not I site and the polyA tail
is CAGCCACGAC. This library was created for the
University of Iowa Mouse Brain Molecular Anatomy Project
(BMAP): "Gene Discovery in the Developing Mouse Nervous
System", supported by National Institute of Mental Health
(NIMH), Hemin Chin, Ph.D., program coordinator."
209 a 177 c 205 g 145 t 1 others
BASE COUNT
ORIGIN
Query Match. 5.9%; Score 334.8; DB 14; Length 737;
Best Local Similarity 69.1%; Pred: No. 7.6e-49;
Matches 505; Conservative 0; Mismatches 213; Indels 13; Gaps 3;
QY 2929 CAGCTGCAAGAGATCAAGCAGCGCTCGCCATCCCGAGCGCAAGATCTCCAAGCA 2988
|||||
Db 20 CAGAGACAAAAGCAGATTCAGACCTCCCGGTGCTCTCCGCAAGAATAGC----- 71
QY 2989 GAAGGAGGAGGAGCGCAAGGATGTCGCATCAGGCAGTACTCCGCCGACGCTTCAAGTT 3048
|||||
Db 72 -AGAGATGAGNAGAGAGAGATCTCCCATCAGAGTTACTCCCTTATGCGCTTCAATTT 130
QY 3049 CTTTCATGGAGCAGCACATAGAGACGTTGATCAATGCTATGCCAGCGCACGATATGCGAA 3108
|||||
Db 131 CTTTCATGGAGCA-CACGTGGGAGATGTGATCAAAACCTACACAGAGGTGAGCGGAG 189
QY 3109 GAATCAGCTGGNAGGAGATGTCACAAGTGGNACTCCCGGATCAGACCCCAATCGAGAT 3168
|||||
Db 190 GCTACAGCTGGAGCAGGAAATGGCCAAAGCTGGGCTCTGTGAGCGCGCAGCAGGACAGAT 249
QY 3169 GAGGAAAATGCTGAACCAAAAGAGAGCAACTACATTCGATTCAAGCGGCCCAAGATGGA 3228
|||||
Db 250 GAGGAGATCTCTTACCAGAGAGGATCTTACTACACCGGCTCAAGAGGCCCAAGATGGA 309
QY 3229 CAAGGAGCATGTTGCTCAAACTGAAGCCCATTTGGAGTGGGTGCATTTTGGCGAGGTTAACGCT 3288
|||||
Db 310 CAAGTCCCATGTTTGTGAAATCAAGACTCTAGGCATCGGTGCGCTTTTGGGAACTGTGCT 369
|||||

```

Qy	3289	GGTGAGCAAAATCGATACCTCGAACCATTGTGTATGCGATGAAACCCCTGCGGAAGCGGA	3348
Db	370	CGCTTGTGAAGCTGGACA---CTCACGCTCTGTACGCCATGAAGACTCTTCAGGAAGAAGGA	426
Qy	3349	CGTTCTCAAGCGGAATCAGGTGGGCACACGTCGAAGCGCGAGAGGGATATCTCTCGCGGAAGC	3408
Db	427	TGTCCTGAACCGGAATCAAGTGGCCCATGTCAAGGCTGAGAGGACATCTCTGGCTGAAGC	486
Qy	3409	CGACAATAAATCGGTGGTGAAGTTGTACTACAGCTCTCCAGGACAAGGATAATCTGTACTTT	3468
Db	487	AGACAATGAGTGGTGGTCAAACTCTACTACTCTCTCCAGGACAAGGACACGCTGTACTTT	546
Qy	3469	TGTGATGAGCTACATACACAGTGGTGTGATCTGATCTGCTCTATCAAACTGGGCACTTTT	3528
Db	547	TGTGATGAGCTACATACACAGCGGGGATATGATGAGCTGCTGATCAGGATGGAGGCTTT	606
Qy	3529	CGAGGAGGAATGCCAGATCTACATCGCGGAGTCCACCTGCGCCCTGGACACCGCTTCA	3588
Db	607	CNCTGAGCACTGGCCCGCTTCTACATGTGACAGTTGACCTTGCCCATTTGAAAGTGTCCTCA	666
Qy	3589	CAAAATGGGCTTCATTACACAGACACATCAAGCCTTCACAACTACTCTATCGATAGGACGG	3648
Db	667	CAAGTGGCTTTATCACCGGACATCAAGCCTTGACACATACTCATCGACCTGGATGG	726
Qy	3649	ACACATAAAGC	3659
Db	727	TCATATTAAAGC	737

RESULT 5	AI517339	644 bp	mRNA	linear	EST 19-APR-2001
LOCUS	GH28053.5	prime GH Drosophila melanogaster head	pot2	Drosophila	
DEFINITION	melanogaster cDNA clone GH28053 5prime similar to U29608: wts				
	P80011739	PID: g903942	SPREMBL: Q24096	mRNA	sequence.
ACCESSION	AI517339				
VERSION	AI517339.1	GI: 4420439			
KEYWORDS	EST.				
SOURCE	fruit fly.				
ORGANISM	Drosophila melanogaster				
	Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota;				
	Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;				
	Ephydroidea; Drosophilidae; Drosophila.				
	1 (bases 1 to 644)				
REFERENCE	Harvey, D., Brokstein, P., Hong, L., Evans-Holm, M., Su, C., Tsang, G.,				
AUTHORS	Lewis, S. and Rubin, G.M.				
TITLE	BDGP/HMMI Drosophila EST Project				
JOURNAL	Unpublished (2001)				
COMMENT	Contact: Stapleton, M.				
	BDGP				
	Lawrence Berkeley National Lab				
	One Cyclotron Rd, Berkeley, CA 94720, USA				
	Fax: 510 486 6798				
	Email: http://www.fruitfly.org/EST_est@fruitfly.berkeley.edu				
	Based upon the presence of vector sequence at both ends, this				
	sequence has been determined to be the complete cDNA insert.				
	Plate: 280 row: E column: 5				
	High quality sequence stop: 561.				
FEATURES	Location/Qualifiers				
source	1..644				
	/organism="Drosophila melanogaster"				
	/db_xref="taxon:7227"				
	/clone="GH28053"				
	/clone_lib="GH Drosophila melanogaster head pot2"				
	/sex="male and female"				
	/dev_stage="adult"				
	/lab_host="DH5-alpha"				
	/note="Organ: head; Vector: pot2; Site: 1: EcoRI; Site 2:				
	pot2; Sized fractionated cDNAs were directly ligated into				
	pot2. Plasmid cDNA library."				
BASE COUNT	145 a 162 c 159 g 178 t				
ORIGIN					

3303 ATACCTCGAACCATTTGTATGCGATGAANAACCTGCGGAAAGCGGACGCTGCTGCAACCGG 3303

High quality sequence stop: 429.

Location/Qualifiers
1 500

... by hydroxyapatite chromatography and used to make this library."

ORGANISM Homo sapiens

Db 241 A---CACTCAGCCCTGTAGGCATGAACACCTTAAGSAAAGAGGATGCTCCTGAACCGGA 297
 QY 3363 ATCAGGTGCACACGTTGAAGGCGGAGAGGATATCTCCGGAAGCGGACATAAATCTGGG 3422
 Db 298 ATCAGGTGCCACGTCAGGCGGAGAGGATATCTCCGGAAGCGGACATAAATCTGGG 357
 QY 3423 TGGTGAAGTTGACTACACGTTCCAGGACAAGATAATCTGACTTTGTGATGGACTACA 3482
 Db 358 TGGTCAAACTACTACTCTCTTCCAGACAAAGACAGCCTGTACTTTGTGATGGACTACA 417
 QY 3483 TACCAGTGGTATGATGTCGCTGCTCATCAAACTGGGCTTTTCGAGGAGGAACTGG 3542
 Db 418 TCCCTGGTGGGACATGATGAGCCTGCTGATCCGGATGAGGATGCTTCCCTGAGCACTGG 477
 QY 3543 CCAGATTTCTACATCGCCGAGGTCACCTGCGCGTGGACAGCGTTCCAAAATAGGGCTTCA 3602
 Db 478 CCGGTTCTACATCGCAGAGCTGACTTTGGCCATTGAGAGTGTCCACAGATGGGCTTCA 537
 QY 3603 TTCACAGACATCAAGCCTGCACACATCTCATCGATAGGAGCGACAT 3654
 Db 538 TCCACCGAGACATCAAGCCTGATGATAATTTTATAGATCTGGATGTCACAT 589

RESULT 9

BI964175
 LOCUS

DEFINITION
 ie65909.y1 Melton Normalized Human Islet 4 N4-HIS 1 Homo sapiens
 cDNA clone IMAGE:5671888 5' similar to TR:Q9Z0W4 Q9Z0W4 LARGE TUMOR
 SUPPRESSOR 1 ; mRNA sequence.

ACCESSION

BI964175

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

Melton, D., Brown, J., Kent, G., Permutt, A., Lee, C., Kaestner, K.,
 Lemishka, I., Scarce, M., Brestelli, J., Gradwohl, G., Clifton, S.,
 Hillier, L., Marra, M., Pape, D., Wylie, T., Martin, J., Blistain, A.,
 Schmitt, A., Theising, B., Ritter, E., Ronko, I., Bennett, J., Cardenas,
 M., Gibbons, M., McCann, R., Cole, R., Tsagarishvili, R., Williams, T.,
 Jackson, Y. and Bowers, Y.
 Endocrine Pancreas Consortium
 Unpublished (2000)

TITLE

JOURNAL

COMMENT

Contact: Douglas Melton, Klaus H. Kaestner, & Hiroshi Inoue
 Endocrine Pancreas Consortium
 Harvard University, Howard Hughes Medical Institute
 Dept of Molecular and Cellular Biology, 7 Divinity Ave, Cambridge,
 MA 02138

Tel: 617-495-1812

Fax: 617-495-8557

Email: dmelton@bioh.harvard.edu

Library was constructed by Dr. Douglas Melton DNA sequencing by:
 Washington University Genome Sequencing Center For information on
 obtaining a clone please contact: Juliana Brown
 (brown@fas.harvard.edu) This sequence now available from the IMAGE
 consortium, for clone orders contact: info@image.llnl.gov
 High quality sequence stop: 432.

FEATURES

source

1. 579

/organism="Homo sapiens"

/db_xref="taxon:9606"

/clone="IMAGE:5671888"

/clone_lib="Melton Normalized Human Islet 4 N4-HIS 1"

/sex="Both"

/tissue_type="Islets of Langerhans"

/dev_stage="Adult"

/lab_host="DH10B"

/note="Organ: Pancreas; Vector: pSPORT1; Site_1: Not 1;

Site_2: Sal 1; Starting library constructed using

SuperScript Plasmid Library kit (Life Technologies). cDNA

made by oligo-dT priming. Size-selected by column
 fractionation; average insert size 1.08 kb. Library was
 amplified once on solid support and plasmid DNA from
 library was prepared. The library DNA was normalized by
 method #4 from Bonaldo, Lennon, and Soares 1996 Genome
 Research 6:791-806; 0.5 microgram single-stranded library
 plasmid DNA was mixed with 5 micrograms PCR product
 representing library inserts and hybridized to an Ecot of
 20. Single-stranded (unhybridized) plasmids were isolated
 by hydroxyapatite chromatography and used to make this
 library."

BASE COUNT 161 a 136 c 163 g 119 t

ORIGIN

Query Match

Best Local Similarity 5.0%; Score 284.4; DB 13; Length 579;

Matches 403; Conservative 0; Mismatches 176; Indels 3; Gaps 1;

QY 3063 ACATAGAGAACGTCATCAAGTCGTATCGCCAGCGCCAGATATCGCAAGATCAAGCTGGAGA 3122
 Db 1 ACGTGGAGAATGTCATCAAAACCTACCACGAGAAGTTAAACCGAGGCTGCAGCTGGAGC 60
 QY 3123 AGGAGATGCACAAAGTGGGACTGCCGGATCAGACCCAAATCGAGATGAGGAAAATGCTGA 3182
 Db 61 AAGAAATGGCCAAAGCTGGACTCTGTGAGCTGAGCAGGAGCAGATGCGGAAGATCCTCT 120
 QY 3183 ACCAAAGGAGAGCAACTACATTCGATTCAAGCGCGCCAGATGACAGAGCATCTTCG 3242
 Db 121 ACCAAGAGAGTCTATTATCAACAGGTTAAGAGGGCCAAAGATGAGCAAGTCTATGTTG 180
 QY 3243 TCAAACTGAAGCCCATTTGGAGTGGGTGCTATTTGGCGAGGTAAACGCTGTGTGACCAAAATCG 3302
 Db 181 TCAAGATCAAAACCTGGGATCGGTGCTTTGGAGAGTGTGCTTCTTGAAGTGG 240
 QY 3303 ATACCTCGAACCATTTGATGGATGAAACCTCGGAAAGCGGACGCTTCTCAAGCGGA 3362
 Db 241 A---CACTCACGCCCTGTACGCCATGAACCCCTAAGGAAAAAGGATGCTCTGAACCGGA 297
 QY 3363 ATCAGTGGCACACCTGAAGCGCGAGAGGATATCTCCGGAAGCGGCACATCACTGGG 3422
 Db 298 ATCAGTGGCCACCTCAAGCGCGAGAGGATATCTCCGGAAGCGGCACATCACTGGG 357
 QY 3423 TGGTGAAGTTGACTACAGCTTCCAGGACAAGATAATCTGTACTTTGTGATGGACTACA 3482
 Db 358 TGGTCAAACTCTACTACTCTTCCAAAGACAAGACAGCCTGTACTTTGTGATGGACTACA 417
 QY 3483 TACAGGTGGTATGATGTCGCTGCTCATCAAACTGGGCAATTTTCGAGGAGGAACTGG 3542
 Db 418 TCCCTGGTGGGACATGATGAGCCTGTGATCCGGATGAGGCTTCTCCCTGAGCACCTGG 477
 QY 3543 CCAGATTTCTACATCGCCGAGGTCACCTCGCGGTGAGCAGGTTCCAAAATGGGCTTCA 3602
 Db 478 CCGGTTCTACATCGCAGAGTGTACTTTGGCCATTGAGAGTGTCCCAAGATGGGCTTCA 537
 QY 3603 TTCACAGACATCAAGCCTGCACACATCACTCATCGATAGG 3644
 Db 538 TCCACCGAGACATCAAGCCTGTATACATTTTCATAGATCTGG 579

RESULT 10

BM801311

LOCUS

DEFINITION
 AGENCOURT_6459219 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5560599
 5', mRNA sequence.

ACCESSION

BM801311

VERSION

KEYWORDS

SOURCE

ORGANISM

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

AUTHORS

NIH-MGC http://mgc.nci.nih.gov/.

957 bp mRNA linear EST 05-MAR-2002

AGENCOURT_6459219 NIH_MGC_88 Homo sapiens cDNA clone IMAGE:5560599

5', mRNA sequence.

BM801311

BM801311.1 GI:19118134

EST

human

Homo sapiens

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

Mammalia; Eutheria; Primates; Catarrhini; Homnidae; Homo.

REFERENCE

1 (bases 1 to 957)

NIH-MGC http://mgc.nci.nih.gov/.

Tissue Procurement: DCID/DIF/GAZDRI

CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
Plate: LLAM9661 row: 1 column: 13
High quality sequence stop: 706.

[illegible]

```

/organism="Homo sapiens"
/db_xref="taxon:9606"
/clone="IMAGE:3885780"
/clone_lib="NIH_MGC_69"
/tissue_type="large cell carcinoma, undifferentiated"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pCMV-SPORT6; Site 1: NotI;
Site 2: SalI; Cloned unidirectionally. Primer: Oligo dt.
Average insert size 1.1 kb. Library constructed by Life
Technologies."
208 a 218 c 271 g 166 t

```

BASE COUNT
ORIGIN

[illegible]

RESULT 14
BG538495

LOCUS BC538495 736 bp mRNA linear EST 03-APR-2001
DEFINITION BC538495.1 GI:13530728 Homo sapiens cDNA clone IMAGE:4691955',
mRNA sequence.
ACCESSION BC538495
VERSION BC538495.1
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominoidea; Homo.
REFERENCE 1 (bases 1 to 736)

AUTHORS
TITLE
JOURNAL
COMMENT

NIH-MGC <http://mgc.nci.nih.gov/>.
National Institutes of Health, Mammalian Gene Collection (MGC)
Unpublished (1999)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: CLONTECH Laboratories, Inc.
cDNA Library Preparation: CLONTECH Laboratories, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA Sequencing by: Incyte Genomics, Inc.
Clone distribution: MGC clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
<http://image.llnl.gov>
plate: LLC1511 row: p column: 04
High quality sequence stop: 570.

FEATURES

/organism="Homo sapiens"
 /db_xref="taxon:9606"
 /clone_image="IMAGE:4691955"
 /clone_lib="NIH_MGC_77"
 /lab_host="DH10B (T1 phage-resistant)"
 /note="Organ: lung; Vector: pDNK-LiB (Clontech); Site_1:
 SfiI (ggcgctcgctgcgc); Site_2: SfiI (ggccattatggcc); 5' and
 3' adaptors were used in cloning as follows: 5' adaptor
 sequence: 5'-CACGGCCATTATGGC-3' and 3' adaptor sequence:
 5'-ATTCTAGAGCCGCGCGGCACATG-DT(30)BN-3' (where B = A,
 C, or G and N = A, C, G, or T). Average insert size 1.9
 kb (range 0.5-4.0 kb). 12/15 colonies contained inserts
 by PCR. This library was enriched for full-length clones
 and was constructed by Clontech Laboratories (Palo Alto,
 CA). Note: this is a NIH_MGC Library."
 183 a 210 c 199 g 144 t

Query Match	4.5%;	Score	255.2;	DB	12;	Length	736;
Best Local Similarity	64.7%;	Pred.	No. 7.1e-35;				
Matches	475;	Conservative	3;	Mismatches	236;	Indels	20;
3533	QY	GAGAACTGGCCAGATCTACATCGCGAGGTCACCTGCGCGTGCAGACAGCTTCACAA	3592				
	Db						
3	Db	GAGCACTGGCCCGGTTCTACATCGCAGAGCTGACITTGCCATTGAGAGTGTCCACAAG	62				
	QY	ATGGGCTTCATTCCACAGAGACATCAAGCCTGACACATACTCATCATAGGAGCGGACAC	3652				
	Db						
63	Db	ATGGGCTTCATCCACCGCAGACATCAAGCCTGATAACATTTGATAGATCTGGATGGTCAAC	122				
	QY	ATAAAGCTCACCGACTTTGGCTGTGCACGGGATTCGGATGGAGCGACACACTCGAAGTAC	3712				
	Db						
123	Db	ATTAAACTCACAGATTTGGGCTCTGCATCGGTTCTCAGGTGGACTCACAATTCCAAATAT	182				
	QY	TACCAAGGAGGAGCCATCTGCAGCAGGACTCGATGAGCGCCCTGGGAGGA-	3765				
	Db						
183	Db	TACCAAGGAGGAGCCATCTGCAGCAGGACAGCATGGAGCCCGACCTCTGGGATGAT	242				
	QY	-----ATACTCCGAGAACCGAACGCCACCGTGTCTGGAGGCGAGGATGCCCGAT	3820				
	Db						
243	Db	GTGTCTAACTGTCTGGTGTGGGACAGGCTGAAGACCCCTAGACAGAGCGCGGAAGCAG	302				
	QY	CACCAAGAGTCTTGCCCACTGCTGTGTGGGACCCCGAACTACATAGCTCCCGAGGTG	3880				
	Db						
303	Db	CACCAAGAGTGGCTGGCACAATTCATGTGTGGGACTTCCAAACTACATCGACCCCGAGGTG	362				
	QY	CTGGAGAGRAGTGGGTACACGACGCTGTCTCGACTACTGTGAGCGTGGCGCTCATCTCTTAY	3940				
	Db						
363	Db	CTCCTCCGCAAGGGTACACTCACTCTGTACTGTGTGGTGGAGTGTCCGAGTGATCTCTTC	422				
	QY	GAGATGCTGGTGGGTCAGCCGCCCTTTCTTGGCCAAACAGTCCGCTGGAAACGCAACAAAG	4000				
	Db						
423	Db	GAGATGCTGGTGGGCGAGCGGCCCTTTTGGCACTACTCCACACAAACCCAGCTGAG	482				
	QY	GTTCATCAAC-TGGGAGAAAACSTGCATATTCCGCCGAGCGCCGAGTATATCCCGGAGGC	4059				
	Db						
483	Db	GTGATCAACTTTGGGAGAACCGCTCCACATTCAGCCCGAGGTGAAGCTGAGCCCTTGAGC	542				
	QY						

```

FEATURES
source
high quality sequence stop: 743.
Location/Qualifiers
1. 824
/organism="Mus musculus"
/strain="CZECH II"
/db_xref="taxon:10090"
/clone="IMAGE:5059023"
/clone_lib="NCI_CGAP_Lu33"
/tissue_type="pooled lung tumors"
/lab_host="DH10B (phage-resistant)"
/notes="Organ: lung; Vector: pT73D-Pac (Pharmacia) with a
modified polylinker; Site_1: NotI; Site_2: EcoRI; 1st
strand cDNA was prepared from mRNA obtained from pooled
lung tumors with a Not I - oligo(dT) primer [5'
TGTTTCCCATCTCAAGTGGAGCGCGCGCTGTGTTTTTTTTTTTTTTT 3'].
Double-stranded cDNA was ligated to Eco RI adaptors
(Pharmacia), digested with Not I and cloned into the Not
I and Eco RI sites of the modified pT73 vector. Library
went through one round of normalization, and was
constructed by Bento Soares and M. Fatima Bonaldo."
BASE COUNT 178 a 232 c 219 g 195 t
ORIGIN
Query Match 4.5%; Score 255.2; DB 13; Length 824;
Best Local Similarity 68.6%; Pred. No. 6.9e-35;
Matches 473; Conservative 2; Mismatches 190; Indels 25; Gaps 8;

```

